

Chapitre IV

Etude des états intermédiaires de repliement.

Un des défis les plus importants dans le domaine du repliement des protéines est de caractériser les structures intermédiaires successives qui jalonnent le chemin de repliement.

Le but principal des études de repliement est de comprendre la liaison qui existe entre cet ensemble séquentiel d'états et la séquence en acides aminés. Ces états peuvent soit se former en conditions d'équilibre soit ne constituer que des espèces de transition apparaissant, en général dans les premières étapes du processus de repliement. Il est très difficile de déterminer les structures intermédiaires lorsqu'il s'agit d'états intermédiaires de transition. En général, il est possible de trouver des conditions expérimentales pour lesquelles ces intermédiaires sont suffisamment stables. Il devient dès lors possible de les étudier, en utilisant les sondes locales structurales classiques. Toutefois de nombreuses protéines donnent lieu à des transitions qui peuvent être décrites par un système à deux états. Ceci est d'autant plus vrai pour les protéines de petite taille à un seul domaine. L'absence d'une conformation stable intermédiaire est une conséquence d'un haut degré de coopérativité observé au cours du processus de repliement. Cette grande coopérativité masque la complexité intrinsèque du processus. Un dépliement coopératif ne signifie pas forcément un équilibre à deux états, parce qu'il faut prendre en compte le fait que chaque courbe de transition, avec une forme sigmoïdale, pourrait être une moyenne de plusieurs transitions très proches.

Beaucoup de progrès ont été faits dans les derniers temps pour mieux comprendre le repliement des petites protéines (Dobson, 1994 ; Fersht, 1995 ; Itzaki, 1995), et il y a

maintenant une forte évidence que les petites protéines aussi peuvent présenter un ou plusieurs états intermédiaires.

Un des moyens classiquement utilisés pour mettre en évidence la présence d'intermédiaires est l'étude combinée des courbes de transition, obtenues au moyen de plusieurs techniques, permettant de corréler les changements thermodynamiques et structuraux.

Ce chapitre est dédié à la mise en évidence et à la caractérisation des états intermédiaires de dénaturation.

Nous avons caractérisé et analysé les régions de transition, à l'équilibre thermodynamique, en fonction de la température et du GdnHCl. Des mesures de microcalorimétrie différentielle, de dichroïsme circulaire, de fluorescence, de diffusion aux petits angles des neutrons et rayons X ont été réalisées pour caractériser, le plus complètement possible, le processus de dépliement. Cela nous a permis de comprendre que le processus de dénaturation, n'est pas un simple équilibre à deux états, que la dénaturation soit induite par la température ou par le chlorure de guanidinium.

4.1 Transition de dénaturation à l'équilibre en fonction de la température

4.1.1. La thermodynamique

La mise en œuvre d'une expérience de calorimétrie est très délicate et nécessite toute une série de tests pour déterminer les conditions optimales de mesure. En effet de nombreux artefacts peuvent affecter les mesures de microcalorimétrie. Tout d'abord, il est important de vérifier que durant le balayage le système est réellement à l'équilibre et qu'il n'existe pas d'effets cinétiques entachant les mesures. D'autre part, il est indispensable de vérifier que les valeurs de dépendent pas de la concentration de la protéine utilisée. Enfin, la réversibilité du processus doit être testée, ceci garantissant la validité des données obtenues.

Une fois établies les conditions de travail, nous avons étudié la réversibilité et les propriétés thermodynamiques qui caractérisent la protéine dans l'intervalle de pH [5,7] en solution d'eau légère.

...la vitesse de chauffe

Entre 15°C et 90°C, pour différents échantillons à pH7, nous avons contrôlé l'indépendance du signal thermodynamique en fonction de la vitesse de chauffe, en faisant varier celle-ci de 60K/h. à 90 K/h. La Fig4. 1, montre la variation de la chaleur spécifique pour la protéine aux deux vitesses. Le signal se montre indépendant de la vitesse de chauffe utilisée, ce qui signifie que pour chaque point le système se trouve à l'équilibre thermodynamique.

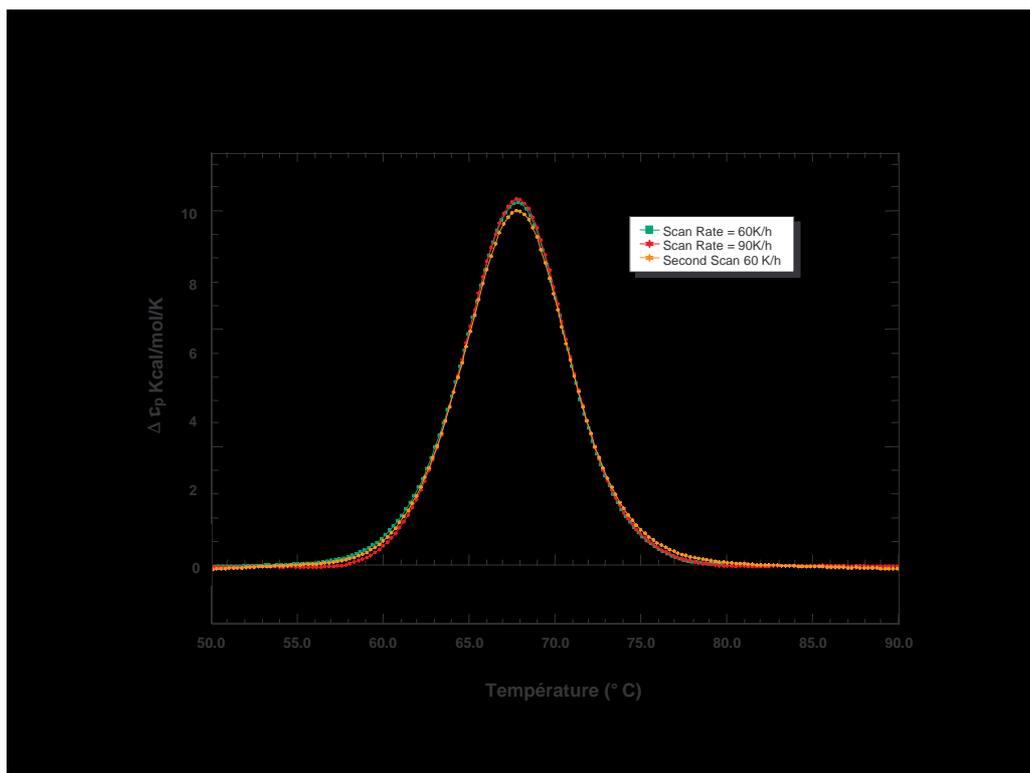


Fig4. 1 Variation de la chaleur spécifique pour la protéine à deux vitesses de chauffe : 60 K/h et 90 K/h. Les courbes rouge et orange représentent des balayages sur le même échantillon.

...les effets de concentration

Il est indispensable de vérifier s'il existe des effets de concentration de protéine. En effet une éventuelle dépendance des transitions en fonction de la concentration constituerait une évidence de l'existence d'intermédiaires intermoléculaires (Filimonov 1996) (apparition de nouveaux pics), ou d'agrégats entachant les valeurs expérimentales. C'est pourquoi nous avons contrôlé la dépendance de la variation de $C_p(T)$ en fonction de la concentration de protéine utilisée. Dans l'intervalle de concentration étudié, entre 0.6 et 3mg/ml, aucune dépendance significative n'a été observée, comme le montre la Fig4. 2 pour l'échantillon à pH7.

Le léger décalage des courbes (T_m varie entre 68.3°C et 68.58) découle essentiellement des variations associées au processus de soustraction de la ligne de base.

Ces contrôles nous permettent de comparer les données obtenues dans différentes conditions expérimentales, comme la fluorescence, les expériences de rayons X aux petits angles, où les concentrations de protéines utilisées sont très différents.

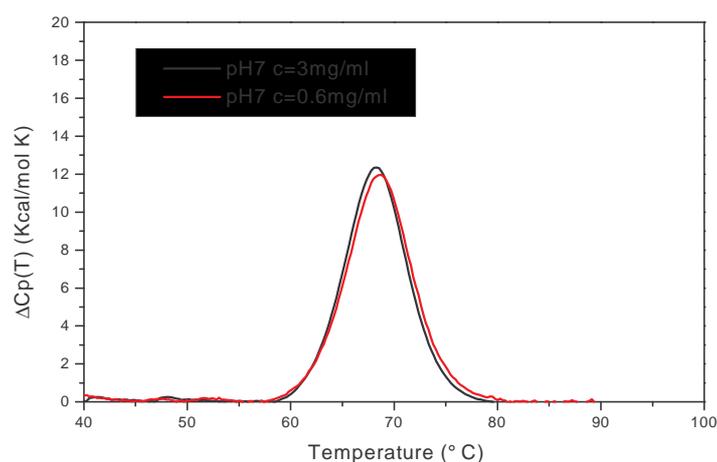


Fig4. 2 Dépendance de la chaleur spécifique de la concentration de la protéine

...variation de la chaleur spécifique en fonction de la température et du pH

Dans les conditions de travail définies ci-dessus, vitesse et concentration, nous avons caractérisé le déploiement de la protéine induit par la température et sa réversibilité aux différents pH 5, 5.5, 6.5, 7, en solution d'eau légère.

L'analyse thermodynamique a été réalisée selon la procédure décrite dans le chapitre II. Cela nous a permis d'obtenir les valeurs de la variation d'enthalpie calorimétrique (ΔH_c) et de celle de Van't Hoff (ΔH_v), la valeur de la température de demi-transition, T_m , et d'estimer le nombre minimal de transitions nécessaires pour ajuster correctement les données.

Toutes les données aux différents pH ont été ajustées d'abord avec un modèle de transition rigoureusement à deux états (enthalpie calorimétrique égale à celle de Van't Hoff), puis avec un modèle qui considère une seule transition mais des valeurs de ΔH_c et ΔH_v différentes, c'est-à-dire une transition *non à deux états*. La Fig4. 3 montre, pour la NCS au pH5 et pH7, un exemple d'ajustement des données avec les deux modèles. On constate que l'ajustement avec un modèle de transition non à deux états est toujours meilleur que celui avec un modèle de transition rigoureusement à deux états.

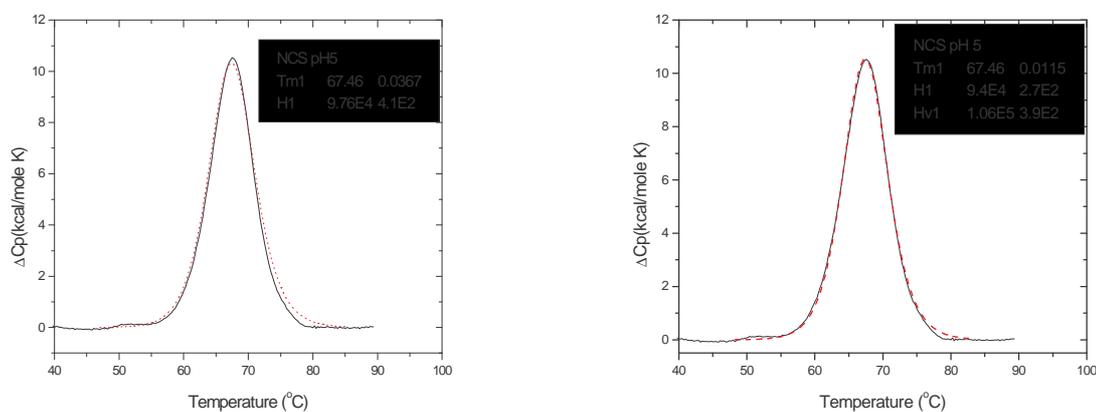


Fig4. 3a Ajustement des données, à pH5, par une transition à deux états et une transition non à deux états.

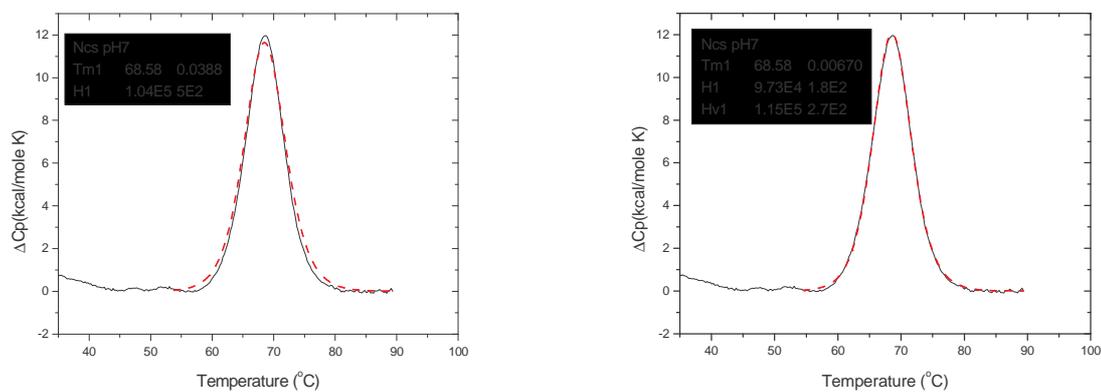


Fig.4.3b Ajustement des données, à pH7, par une transition à deux états et une transition non à deux états

Les résultats dérivés de l'analyse de chaque échantillon, sont résumés dans le tableau 4.1. Tous les paramètres thermodynamiques obtenus en fonction du pH sont, dans la barre d'erreur expérimentale, équivalents, c'est-à-dire qu'aucune variation significative n'a été observée. Ceci correspond au fait que la valeur de la variation de la chaleur spécifique entre l'état natif et l'état dénaturé est très faible dans cet intervalle de pH. Ces résultats montrent, comme les résultats de diffusion de la lumière, que la NCS n'est pas sensible au pH dans la plage explorée.

Toutefois, il est important de noter que des problèmes apparaissent si le pH est supérieur à 7.5 (à cause de la présence des ponts disulfure) ou inférieur à 4 (proximité du point isoélectrique).

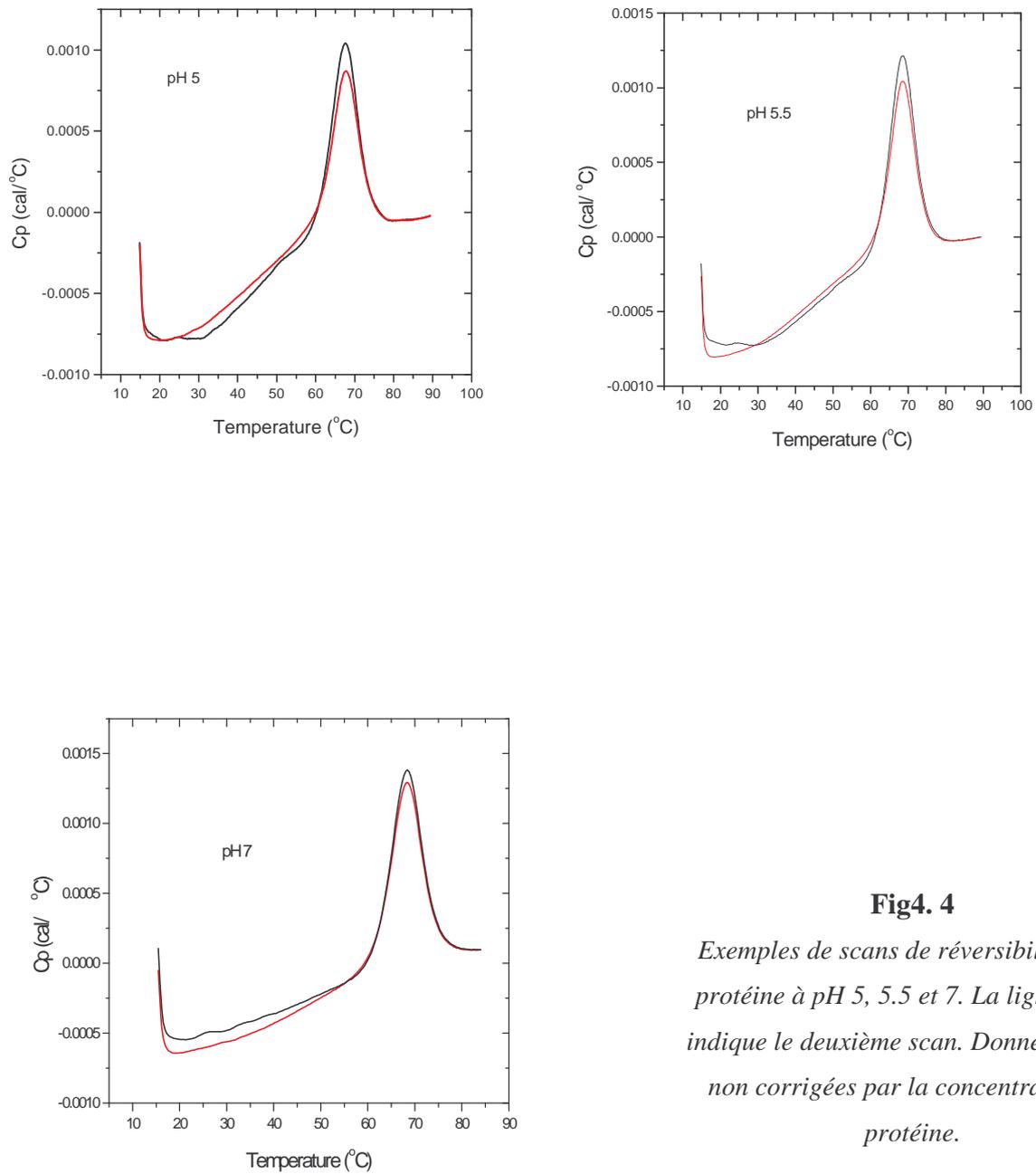
Tableau 4.1

PH	T_m (°C)	ΔH_c (Kcal/mol)	ΔH_v (Kcal/mol)
7 en H ₂ O	68.58	97.3	115
6.5 en H ₂ O	68.58	92.8	113
5.5 en H ₂ O	68.50	94	107
5 en H ₂ O	67.46	94	106

....la réversibilité de la courbe de transition de dénaturation.

Pour étudier la réversibilité, chaque échantillon a été chauffé et refroidi, dans la même cellule deux ou trois fois de suite. Dans la Fig4. 4 nous présentons les scans consécutifs sur la protéine à pH5, pH5.5, pH7. Le même comportement a été obtenu pour les autres échantillons.

Les spectres coïncident parfaitement dans la forme, les transitions étant symétriques et avec la même température de demi-transition. La conservation d'une forme symétrique est très importante car elle indique que nous ne sommes pas en présence d'espèces oligomériques pouvant se former pendant le refroidissement. En revanche, les amplitudes maximales des spectres ne coïncident qu'à 90% ou 95% environ.

**Fig4. 4**

Exemples de scans de réversibilité de la protéine à pH 5, 5.5 et 7. La ligne rouge indique le deuxième scan. Données brutes non corrigées par la concentration de protéine.

Nous pouvons en déduire qu'un petit pourcentage de protéines ne se replie pas, ou se replie partiellement, mais sans conduire à la formation d'agrégats, parce que, dans ce dernier cas, la forme du pic de transition serait asymétrique et la valeur de T_m serait changée. En regardant la fig (4.4) nous pouvons remarquer que dans la région où la protéine est complètement dénaturée, les courbes se superposent, ce qui veut dire que la quantité de protéine dépliée est la même dans les deux expériences. Dans la première partie du spectre et dans la zone de la transition les courbes sont décalées ; cette variation est sûrement due au fait que les concentrations de protéine native dans le premier scan (c1) et dans le deuxième scan (c2) sont faiblement différentes ($c1 > c2$).

Ces études montrent que le processus de repliement est globalement conservé dans l'intervalle de pH compris entre 5 et 7. Dans cet intervalle, la *transition de dénaturation est réversible à 95%*, caractéristique fondamentale pour les études de repliement. D'autre part, l'ensemble des données obtenues suggèrent fortement que *la transition de dénaturation semble plus complexe qu'un simple processus deux états*.

4.1.2 Variation de la structure secondaire et tertiaire en fonction de la température.

Les courbes de dénaturation, qui représentent l'évolution de la structure secondaire et tertiaire de la protéine, ont été établies en suivant la variation de l'ellipticité à 223 nm et de la longueur d'onde d'émission maximale de fluorescence dans l'intervalle de température compris entre 20°C et 80°C.

La Fig4. 5 montre la variation de l'ellipticité de la protéine, en H₂O, en fonction de la température. Comme nous l'avons déjà vu la NCS présente un spectre particulier de dichroïsme, et l'information que nous suivons représente la variation des contraintes internes de la protéine. La courbe, obtenue par la variation de l'intensité à 223 nm en fonction de la température, a une allure globalement symétrique, avec un T_m de 63.7°C, et une perte totale de signal dichroïque à la température de 70°C.

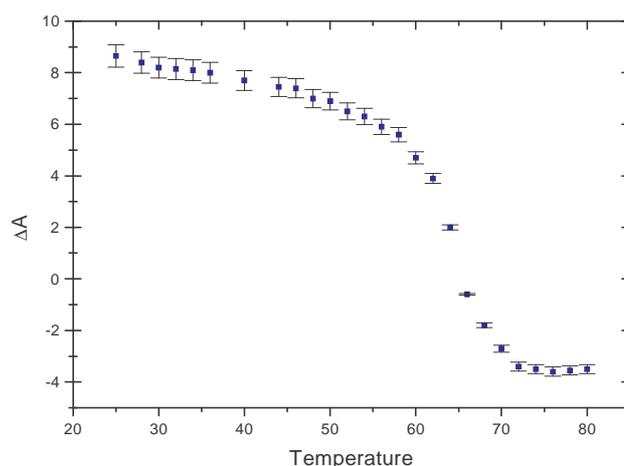


Fig4. 5 Transition de dénaturation de la NCS par la température suivie par la variation d'ellipticité en eau légère.

Les courbes obtenues par fluorescence des tryptophanes, représentées en Fig4. 6 sont bien décrites par une simple transition, soit en eau légère soit en eau lourde. Comme nous le verrons aussi pour les courbes décrites en fonction du chlorure de guanidinium, les longueurs d'émission maximales sont très élevées, en raison de la forte exposition des chromophores à l'eau.

Les deux transitions, en H₂O et en D₂O, sont décalées en température, avec un T_m de 68°C et 73°C respectivement. Ces valeurs sont complètement comparables avec les valeurs déterminées par les études de calorimétrie (voir section 3.2.1). Comme attendu, la protéine, est plus stable dans l'eau lourde que dans l'eau légère. La protéine conserve un signal du type natif jusqu'à la température de 60°C en H₂O et 65°C en eau lourde, et se dénature complètement à 78°C et 81°C respectivement. Ces températures étant très élevées, nous n'avons pas pu caractériser complètement les régions du plateau de dénaturation.

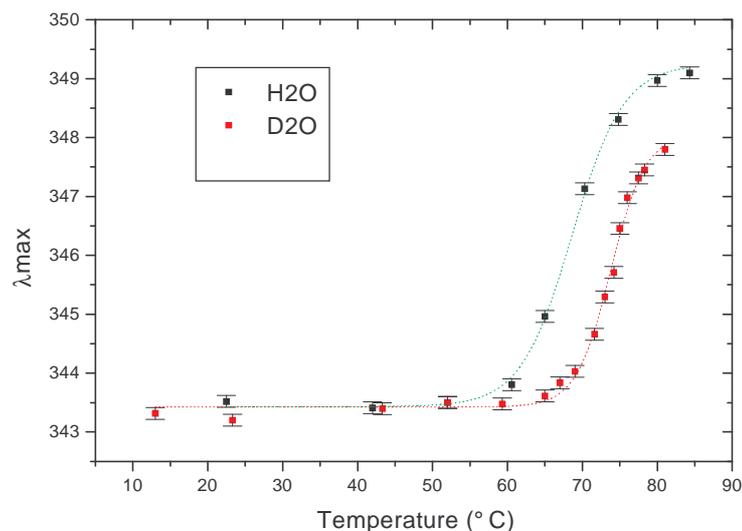


Fig4. 6 Transition de dénaturation établies en fluorescence en fonction la température, en eau légère et en eau lourde.

Les courbes de transition ne sont pas seulement décalées en température mais aussi en amplitude. La transition en D₂O présente toujours des longueurs d'onde maximales plus basses, et ceci même pour la forme complètement dénaturée.

Ce résultat n'est pas surprenant dans la mesure où la valeur caractéristique de 349 nm du tryptophane totalement exposé au solvant se réfère à l'acide aminé en H₂O. Le décalage obtenu dans l'eau lourde découle du fait que le signal de fluorescence dépend des interactions directes et indirectes du fluorophore avec le milieu environnant (Bergethon, 1998).

L'analyse thermodynamique des courbes de transition de dénaturation suivies en CD et fluorescence a été effectuée en utilisant l'approximation d'une transition à deux états. Ceci nous a permis de déterminer le point de demi transition T_m , mais aussi la variation d'enthalpie ΔH_m associée. L'ajustement de l'équation (1.6) aux courbes de fluorescence des tryptophanes a pu être correctement réalisé. Nous avons alors estimé une valeur d'enthalpie de 66.5 ± 1 kcal/mol associée au signal de fluorescence détecté en H₂O, et une valeur de 81 ± 1 kcal/mol pour la transition en D₂O.

En revanche l'ajustement a été plus difficile pour la courbe de dichroïsme circulaire.

Les valeurs d'enthalpie obtenues par fluorescence s'écartent de celles déterminées par la microcalorimétrie dans les mêmes conditions de solvant. Il est évident que si l'équilibre de dénaturation en température de la protéine est vraiment un équilibre à deux états, les résultats expérimentaux comme la variation d'enthalpie, et les points de demi-transition etc..., qui sont extrapolés des courbes de transitions, doivent être indépendants de la sonde utilisée et donc être équivalents aux barres d'erreur expérimentales près. Ceci nous suggère que si, les courbes spectroscopiques sont bien décrites, en première approximation, par une transition à deux états, elles masquent des informations très importantes, que nous allons mettre en évidence dans le prochain paragraphe à l'aide de la diffusion aux petits angles des rayons X.

4.1.3. Changement de conformation suivie par diffusion aux petits angles des rayons X.

Les expériences de diffusion aux petits angles des rayons X nous ont permis de déterminer la variation du rayon de giration de la NCS quand on augmente la température de 20°C à 80°C. L'analyse des courbes de diffusion dans un intervalle de vecteur de diffusion Q compris entre 0 et 0.3 \AA^{-1} , nous a permis aussi de suivre la variation de la forme au cours de la dénaturation, à l'aide de la fonction de distribution des distances et de la représentation des spectres en $Q^2I(Q)$, dite représentation de Kratky.

4.1.3.1 Variation de la taille

Le premier problème à résoudre a été le suivant : quel type d'analyse des courbes de diffusion doit-on utiliser pour la détermination du rayon de giration, dans la région de la transition? Nous avons vu dans la section 1.3.4 que le type d'analyse est dépendant de la structure de la protéine. Il n'est pas le même suivant que la protéine est globulaire compacte, ou dépliée : une bonne description des spectres est donnée pour la protéine native par l'approximation de Guinier, et pour la protéine complètement dénaturée par la théorie des polymères avec la loi de Debye. Dans la zone de transition, le problème est plus compliqué puisque la solution de protéines est un mélange de plusieurs populations : native, totalement dépliée, et éventuellement dans un état intermédiaire. Dans ce cas, l'intensité mesurée est donnée par l'équation (1.49) et le rayon de giration "moyen" théorique par l'équation (1.53).

Afin de déterminer quel type d'analyse est le plus approprié (Guinier ou Debye) pour obtenir une mesure correcte du rayon de giration "moyen", nous avons simulé l'intensité

théoriquement obtenue dans le cas le plus simple d'une transition à 2 états, natif et complètement déplié. Si la dilution est infinie, cette intensité théorique se met sous la forme :

$$I_T(Q) = f_N I_N(Q) + (1 - f_N) I_D(Q) \quad (4. 1)$$

où f_N est la fraction de protéine native.

$I_N(Q)$ est l'intensité théorique diffusée par la protéine native, assimilée à une sphère de rayon R , et donc de rayon de giration $R_N = (5/3)^{1/2} R$. Elle a l'expression suivante :

$$I_N(Q) = \left(3 \frac{\sin QR - QR \cos QR}{Q^3 R^3} \right)^2$$

Pour $I_D(Q)$, nous avons pris l'intensité diffusée par une chaîne de Kratky Porod, et représentée par l'expression de Sharp et Bloomfield (équation 1.55).

L'intensité théorique $I_T(Q)$ a été calculée pour différentes valeurs de f_N , en prenant pour le rayon de la protéine native $R = 18 \text{ \AA}$ (ce qui correspond à un rayon de giration R_N de 14 \AA), et en affectant à la protéine complètement dépliée les valeurs des paramètres L , b et $R_{g,c}$ obtenues expérimentalement et données dans la section 3.2.3.1 (qui conduisent à un rayon de giration $R_D = 26.2 \text{ \AA}$). Nous avons ensuite extrait un rayon de giration en ajustant les courbes simulées $I_T(Q)$, soit à l'approximation de Guinier, soit à la loi de Debye, dans la même zone de vecteur de diffusion : $0.02 \text{ \AA}^{-1} < Q < 0.055 \text{ \AA}^{-1}$. Cette zone d'ajustement est raisonnable puisque sa valeur maximale $Q_{\max} = 0.055 \text{ \AA}^{-1}$ correspond à $Q_{\max} R_g < 1.4$. La comparaison des rayons de giration R_{guinier} et R_{debye} ainsi obtenus au rayon de giration "moyen" théorique du mélange :

$$R_{th} = \sqrt{f_N R_N^2 + (1 - f_N) R_D^2} \quad (4. 2)$$

est présentée sur la Fig4. 7 pour différentes fractions f_N de protéine native.

On observe sur cette figure que, pour la zone d'ajustement considérée, la loi de Debye constitue clairement une bonne approximation dès que la protéine commence à se déplier et dans toute la région de transition. Par contre, l'approximation de Guinier conduit à des valeurs de R_g trop faibles. On pourrait objecter que la loi de Guinier donnerait des valeurs plus proches de la valeur théorique si l'on réduisait encore la zone d'ajustement, mais cela est techniquement difficile puisqu'il faut que cette zone contienne un nombre suffisant de points de mesures. Nous avons vérifié par ailleurs que, si l'on prend une zone d'ajustement plus grande que celle correspondant à $Q_{\max} R_g < 1.4$, l'écart entre le rayon "moyen" théorique R_{th} et le rayon obtenu à partir de la loi de Debye augmente, rendant ainsi l'utilisation de la loi de Debye moins satisfaisante. En résumé, il est préférable d'utiliser l'approximation de Debye

pour extraire le rayon de giration "moyen" d'un mélange de populations, native et dénaturées, à condition de faire l'ajustement dans une zone en Q suffisamment réduite ($Q_{\max}R_g < 1.4$).

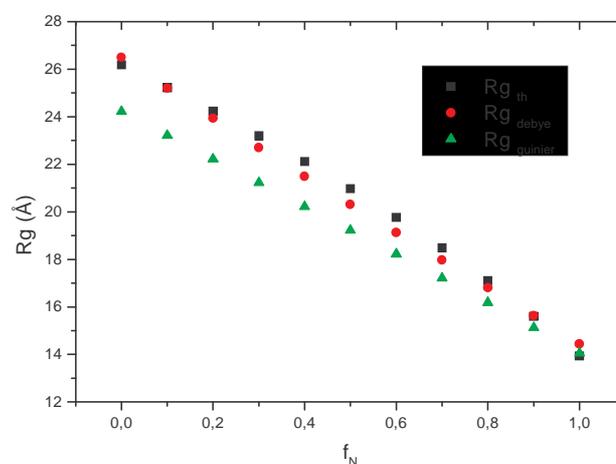


Fig4. 7 Variation, avec la fraction f_N de protéine native, du rayon de giration théorique R_{th} , et des rayons de giration $R_{guinier}$ et R_{debye} déduits des ajustements des courbes de diffusion théoriques $I_T(Q)$ aux lois de Guinier et Debye respectivement, dans le cas d'une transition à deux états.

Nous avons tenu compte de ces considérations pour déterminer le plus précisément possible le rayon de giration en fonction de la température. Les courbes de diffusion, mesurées pour les températures inférieures à 65°C ont été analysées avec l'approximation de Guinier, dans le domaine en Q : $0.02 \text{ \AA}^{-1} < Q < 0.055 \text{ \AA}^{-1}$. Par contre, nous avons utilisé la loi de Debye pour toutes les températures supérieures, en faisant l'ajustement dans le même domaine en Q . Les résultats des ajustements des différentes courbes de diffusion sont présentés sur la Fig4. 8. Notons sur cette figure l'allure de $I(Q)$ pour les plus petites valeurs de Q , qui indique une absence presque totale d'agrégats, quelle que soit la température. Il faut également mentionner que ces mesures ont été effectuées pour une concentration de NCS de 5mg/ml. Sur les courbes présentées ci-dessus, nous n'avons pas fait la correction correspondant aux effets d'interaction. Rappelons que cette correction est très faible dans les conditions où nous avons travaillé, elle est même nulle quand la protéine est dépliée, puisque nous avons vu que le second coefficient du Viriel est alors voisin de zéro.

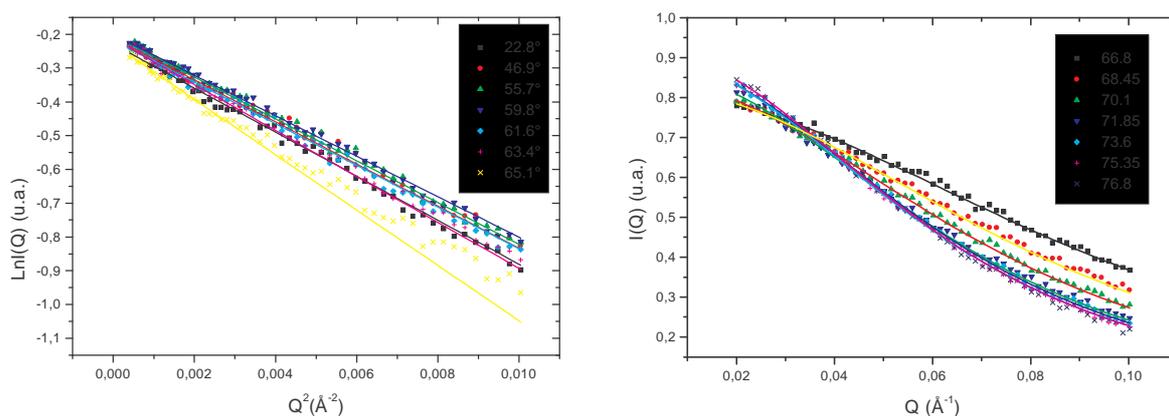


Fig4. 8 Ajustement des différentes courbes de diffusion par la l'approximation de Guinier(gauche) et par la loi de Debye (droite).

Sur la Fig4. 9 nous présentons les valeurs du carré du rayon de giration R_g^2 en fonction de la température, issues des ajustements précédents. La grandeur pertinente pour décrire la transition est en effet R_g^2 et non R_g , puisque c'est R_g^2 qui varie linéairement en fonction des fractions des différentes espèces en présence. A titre de comparaison, nous montrons également sur la figure 4.9 les valeurs du carré du rayon de giration obtenues en utilisant l'approximation de Guinier quelle que soit la valeur de la température.

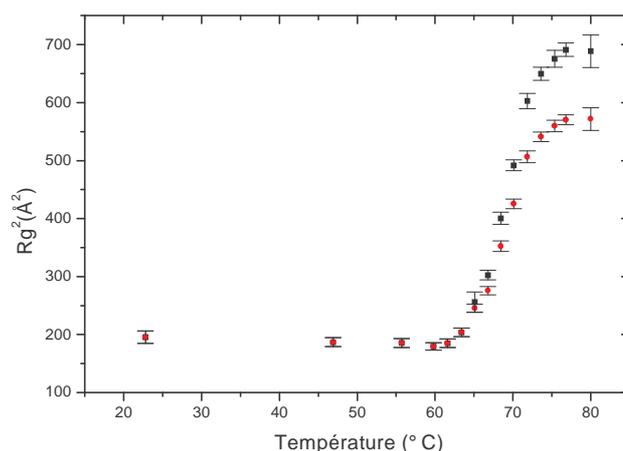


Fig4. 9 Transition de dénaturation en température, suivie par la variation du carré du rayon de giration, obtenue en ajustant les courbes de diffusion mesurées à 5mg/ml, soit à la loi de Debye(en noir), soit à l' approximation de Guinier(en rouge).

Comme il était attendu, ces valeurs sont plus sensiblement plus faibles que celle obtenues par Debye dès que la protéine se déplie.

On observe que la protéine garde un rayon de giration de type natif jusqu'à 65°C environ. On note même une légère décroissance du rayon de giration quand la température passe de 22 à 60°C. Ceci provient du fait que les mesures ci-dessus ont été effectuées à 5mg/ml, et que les effets de concentration sont un peu plus importants à 60°C qu'à 22°C. La légère décroissance de R_g ($c=5\text{mg/ml}$) correspond à une petite augmentation du second coefficient du Viriel A_2 , mais le rayon de giration extrapolé à concentration nulle R_g ($c=0$) est en fait constant (voir la section 3.1.1.2).

A partir de 65°C, le rayon de giration commence à augmenter. La transition de la forme globulaire à la chaîne dépliée se produit sur un intervalle d'une dizaine de degrés. Au-delà de 76°C, R_g présente à nouveau un plateau.

La courbe de transition du carré du rayon de giration en fonction de la température présente, comme celle de la fluorescence, une allure symétrique, et peut paraître typique d'une transition à deux états. L'analyse thermodynamique de cette courbe avec le modèle de Pace nous a permis de déterminer la valeur de la température de demi-transition et la variation d'enthalpie associée. L'ajustement de la courbe à l'équation (1.6) conduit à $T_m=69\pm 1^\circ\text{C}$, et $\Delta H_m=132\pm 1$ Kcal/mol.

Nous sommes amenés à faire la même remarque que dans la section précédente : bien que le point de demi-transition soit équivalent à ceux déterminés précédemment par fluorescence ou microcalorimétrie, il n'en est pas de même pour la valeur du ΔH_m , qui diffère de celle obtenue en fluorescence et de celle obtenue en microcalorimétrie. Un modèle à deux états n'est donc pas suffisant pour décrire complètement la transition.

L'intensité à $Q=0$

La seconde quantité extraite de l'analyse des courbes de diffusion dans la région des petites valeurs de Q est la valeur de l'intensité à l'origine, $I(0,c)$. Bien que nos expériences de

rayons X n'aient fourni qu'une détermination relative, et non absolue, des $I(0,c)$, il est possible d'en déduire deux types d'informations intéressantes. Le premier concerne l'évolution de la masse M des particules diffusantes, et permet de contrôler l'absence de formation d'agrégats au cours de la dénaturation. Le second concerne l'évolution du volume spécifique partiel de la protéine quand elle se déplie.

La Fig4. 10.présente la variation en température de $I(0,c)$ pour la concentration de protéine $c=5\text{mg/ml}$. Pour tirer de cette courbe des informations quantitatives, il est nécessaire de corriger les valeurs de $I(0,c)$ des effets d'interaction. Nous avons pour cela utilisé l'expression (1.69) qui permet de déduire $I(0,0)$ de $I(0,c)$, moyennant la connaissance du second coefficient du Viriel A_2 . Les valeurs de A_2 avaient été déterminées juste avant ($A_2 \approx 5 \cdot 10^{-4} \text{ cm}^3 \text{ mol g}^{-2}$) et juste après la transition ($A_2 \approx 0$) (voir la section 3.1.1.2). Nous avons ensuite fait l'hypothèse que le second coefficient du Viriel A_2 suit une variation analogue à celle du carré du rayon de giration quand la température passe de 65 à 75°C. Il est bien clair qu'il s'agit ici d'une approximation, mais qu'elle est suffisante pour connaître la variation de $I(0,0)$ entre les états initial et final de la protéine. Les valeurs de $I(0,0)$ ainsi obtenues sont montrées sur la Fig4. 10.

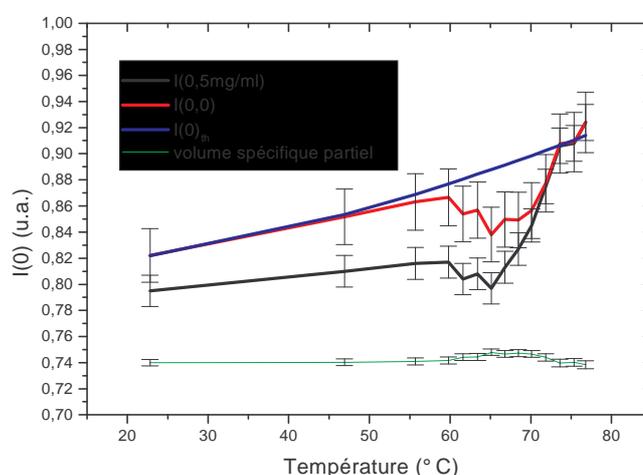


Fig4. 10 Variation avec la température de l'intensité à l'origine $I(0,c=5\text{mg/ml})$ déterminée expérimentalement, de l'intensité $I(0,0)$ extrapolée à concentration nulle, et de l'intensité théorique $I_{th}(0,0)$ que l'on aurait en l'absence de variation de volume spécifique partiel.

Variation avec la température du volume spécifique partiel, estimée à partir de $I(0,0)$ en utilisant l'équation 1.48.

Pour comprendre le comportement en température de $I(0,0)$, il faut utiliser l'expression (1.48) qui relie $I(0,0)$ au volume spécifique partiel de la protéine et à la densité du solvant. Cette expression permet de calculer la variation en température de l'intensité à l'origine $I(0,0)_{th}$ que l'on aurait en l'absence de variation de volume spécifique partiel, et qui serait due uniquement à la variation de la densité du solvant ρ_s quand la température augmente (notons que la diminution de ρ_s avec la température entraîne une diminution de la concentration de protéine, dont il faut aussi tenir compte). La courbe de $I(0,0)_{th}$ est présentée sur la Fig4. 10. On observe que les points expérimentaux de $I(0,0)$ ne diffèrent que de 5 à 6% au maximum de la courbe théorique $I(0,0)_{th}$, et seulement au voisinage du point de demi-transition. C'est un résultat très satisfaisant, dans la mesure où il indique sans ambiguïté que nous n'avons eu aucun problème d'agrégation. On peut également chercher à déduire la variation en température du volume spécifique partiel, à partir des valeurs expérimentales de $I(0,0)$, en utilisant encore l'expression (1.48), et en prenant la valeur de $0.74 \text{ cm}^3\text{g}^{-1}$ pour le volume spécifique partiel de la protéine native. Les résultats sont donnés sur la Fig4. 10. Ils sont évidemment à considérer avec beaucoup de recul, puisque d'une part les mesures d'intensité comportent une certaine imprécision, et d'autre part l'obtention du volume spécifique partiel a nécessité plusieurs hypothèses, notamment concernant le comportement de A_2 , et la diminution en température de la concentration de protéine. Il faut également souligner qu'il s'agit d'un volume spécifique "moyenné" sur les différentes populations en présence. Ces résultats sembleraient toutefois suggérer que le volume spécifique partiel de la protéine subit une légère augmentation au cours de la transition de dénaturation, puis retrouve sa "valeur native" quand la protéine est complètement dépliée. Un commentaire s'impose: quand la protéine est dépliée, sa surface accessible au solvant est plus grande que quand elle est native, et donc sa couche d'hydratation a un volume plus grand. Si cette couche était aussi structurée que pour la protéine native, on devrait alors s'attendre à un volume spécifique partiel plus faible pour la protéine dénaturée que pour la protéine native, ce qui ne semble pas être le cas. On peut alors se demander si nos résultats n'indiquent pas que la couche d'hydratation de la protéine dénaturée par chauffage est moins structurée que celle de la protéine native. Ce point mériterait d'être ultérieurement précisé.

4.1.3.2 Fonction de distribution des distances

L'analyse précédente concerne la région des petites valeurs de Q des courbes de diffusion, et permet essentiellement d'obtenir des informations sur la compacité de la protéine. Des informations supplémentaires sur la conformation de la particule peuvent être obtenues en utilisant l'ensemble de la courbe de diffusion, et en calculant la fonction de distribution des distances $P(r)$, qui est la transformée de Fourier inverse de l'intensité diffusée $I(Q)$:

$$P(r) = \frac{r^2}{2\pi^2} \int_0^\infty Q^2 I(Q) \frac{\sin(Qr)}{Qr} dQ$$

où r est la distance entre deux points diffuseurs (voir la Fig4. 11). $P(r)$ représente le nombre d'éléments de volume de la particule distants de r , en d'autres termes l'histogramme des distances de la particule. $P(r)$ s'annule pour $r < D_{\max}$, où D_{\max} est la dimension maximale de la particule.

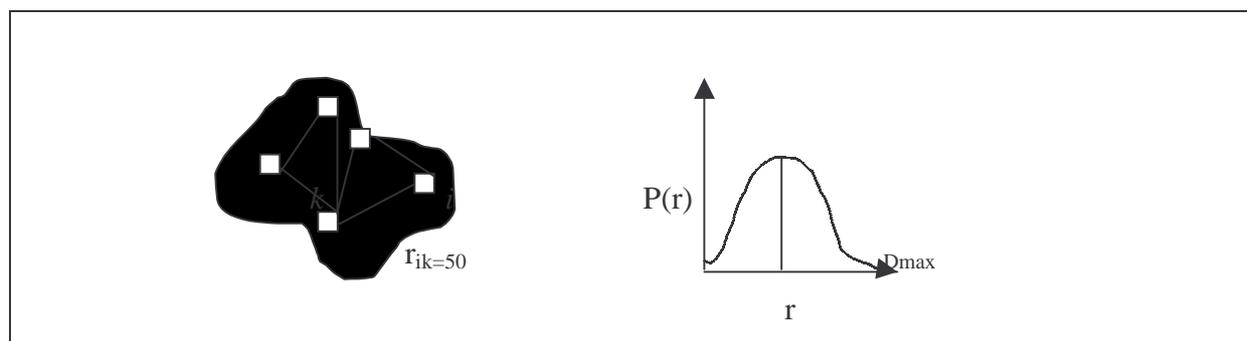


Fig4. 11 Représentation schématique d'une fonction de distribution des distances $P(r)$.

La fonction $P(r)$ contient les mêmes informations que l'intensité diffusée $I(Q)$, mais, contrairement à l'intensité, elle donne accès directement aux distances interatomiques à l'intérieur de la particule, et donc à sa structure. On peut extraire en particulier une quantité importante : l'extension maximale de la particule, D_{\max} . On peut également obtenir le rayon de giration à partir de $P(r)$, en utilisant la relation :

$$R_g^2 = \frac{\int r^2 P(r) dr}{2 \int P(r) dr}$$

Cette méthode de détermination de R_g est intéressante parce qu'elle utilise l'ensemble de la courbe de diffusion $I(Q)$, alors que la détermination du R_g à partir des lois de Guinier ou Debye ne fait appel qu'à la partie des petites valeurs de Q de cette courbe. Elle sera donc moins sensible aux effets qui peuvent perturber la courbe $I(Q)$ à petits Q , comme les interactions entre particules. Son second intérêt est qu'elle ne fait appel à aucune approximation (Guinier ou Debye). Elle a toutefois pour inconvénient de fournir des valeurs de R_g avec une barre d'erreur non négligeable, et il est finalement meilleur de coupler son utilisation à celle des approximations de Guinier ou Debye.

Tout le problème réside dans la restauration de la fonction $P(r)$ à partir de l'intensité diffusée $I(Q)$. Différentes méthodes, dites de "transformée indirecte", existent pour résoudre ce problème. Le programme que nous avons utilisé est celui développé par Svergun (programme GNOM, avec les contraintes standard) (Svergun et al., 1988; Svergun, 1992). Les courbes $P(r)$ ainsi obtenues pour les différentes températures sont présentées sur la Fig4. 12. On observe que la forme de la fonction $P(r)$, ainsi que la position de son pic, ne varient pratiquement pas jusqu'à 66°C. A partir de cette température, l'extension maximale de la particule commence à augmenter. Par contre, la position du pic varie peu jusqu'au point de demi-transition (69°C), ce qui est compatible avec le fait qu'il subsiste jusque là une population native. Ce n'est qu'au delà de ce point que la forme et la position du pic de $P(r)$ changent nettement, pour conduire finalement à 75°C au profil caractéristique d'une chaîne dépliée.

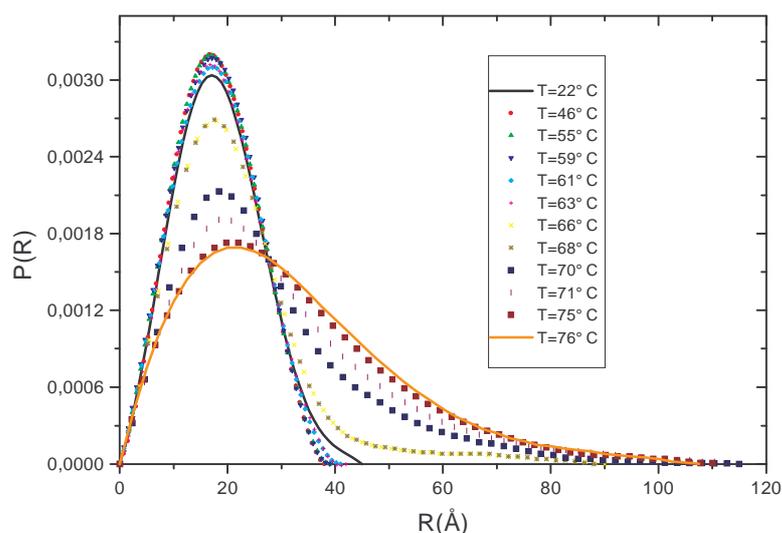


Fig4. 12 Fonction de distribution des distances $P(r)$, de la NCS, pour les différentes températures.

Les rayons de giration calculés à partir des courbes $P(r)$ sont donnés sur la Fig4. 13, en comparaison avec les rayons de giration déterminés dans la région des petites valeurs de Q , à l'aide de la loi de Debye. On constate que l'accord est excellent, ce qui nous conforte dans l'idée que la loi de Debye constitue une excellente approximation quand on veut déterminer le rayon de giration "moyen" d'un mélange contenant des espèces non globulaires.

La courbe 4.13 montre également la variation de D_{\max} en fonction de la température. Bien que la détermination de D_{\max} ne soit pas très précise, on observe que sa variation avec la température est plus rapide que celle du rayon de giration. Jusqu'à 65°C , D_{\max} reste voisin de 40\AA , puis elle atteint rapidement sa valeur maximale de $107\pm 6\text{\AA}$; le plateau au voisinage de cette valeur s'étend à partir de 70°C .

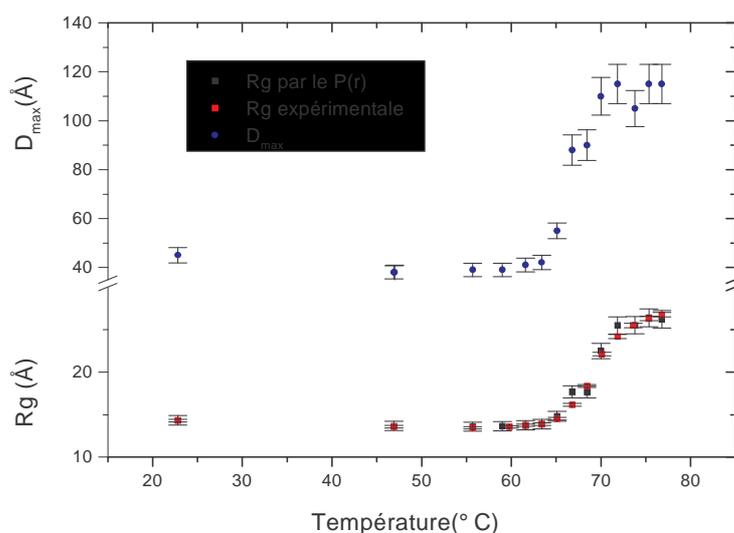


Fig4. 13 Variation, en fonction de la température, du rayon de giration obtenu à partir de la courbe de distribution des distances, avec en comparaison le rayon de giration déterminé à partir de l'ajustement à la loi de Debye.

Variation, en fonction de la température, de la dimension maximale D_{\max} de la protéine.

La compréhension du comportement de D_{\max} requiert un commentaire. L'extension maximale de la protéine, déterminée à partir de $P(r)$, n'est pas une grandeur qui permet, comme le carré du rayon de giration "moyen", de suivre la courbe de transition, en ce sens

qu'elle ne varie pas linéairement en fonction des fractions de populations des diverses espèces. Le D_{\max} ainsi déterminé représente l'extension de l'espèce la plus étendue présente dans la solution. Dans le cas d'une transition purement à 2 états, natif et complètement déplié, on s'attend donc à ce que D_{\max} ait un comportement en "marche d'escalier" en fonction de la température, passant de manière abrupte de la valeur de la protéine native à celle de la protéine dénaturée. Le comportement observé dans le cas de la NCS n'est pas tout à fait en "marche d'escalier", et le fait que quelques valeurs de D_{\max} (deux points de mesure vers 90 Å) soient intermédiaires entre 40 et 107 Å laisse déjà supposer que des espèces non natives et non complètement dépliées sont présentes.

4.1.3.3 Représentation de Kratky des courbes de diffusion

Dans le but d'avoir des informations supplémentaires sur l'évolution de la forme globale de la protéine, les courbes de diffusion ont été tracées en représentation de Kratky, $Q^2I(Q)$ en fonction de Q . Cette représentation est souvent utilisée dans les études de dénaturation, parce que le profil obtenu est clairement différent suivant que la protéine est dans une forme globulaire ou dans une forme dépliée (voir la section 1.3.4.4).

Pour pouvoir comparer les différentes courbes, il est nécessaire de normaliser chaque spectre $I(Q,c)$ par la valeur de l'intensité à l'origine $I(0,c)$ correspondante. Il faut en effet s'affranchir du terme de contraste, pour ne conserver que le facteur de forme. Les courbes normalisées $Q^2I(Q,c)/I(0,c)$ sont représentées sur la Fig4. 14 Le spectre à 22°C a une forme de "cloche", typique d'une protéine native. A partir de 65°C, l'amplitude du pic commence à décroître, pour conduire finalement, à 76.8°C, à une courbe présentant un plateau pour $Q > 0.1 \text{ \AA}^{-1}$, et qui caractérise la protéine complètement dépliée.

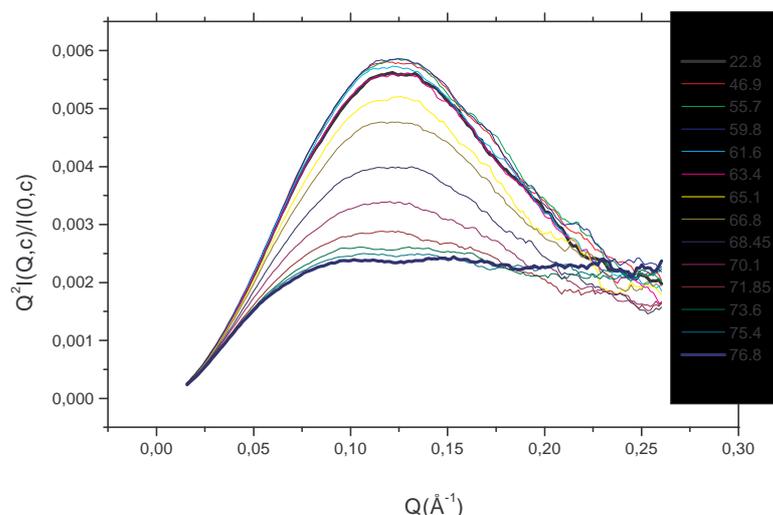


Fig4. 14 Représentation de Kratky des courbes de diffusion $I(Q,c)$ normalisées à $I(0,c)$, avec $c=5\text{mg/ml}$, obtenues pour les différentes températures. Les courbes ont été lissées afin de s'affranchir du bruit expérimental. La procédure de lissage consiste, pour chaque valeur de Q , à faire une moyenne de $I(Q)$ sur les 5 points expérimentaux de part et d'autre du point considéré. Il a été soigneusement vérifié que ce lissage a pour unique conséquence de supprimer le bruit.

Le fait remarquable mis en évidence sur cette figure est que, dans la région de transition, c'est-à-dire entre 65 et 75°C, **les différentes courbes de diffusion ne se croisent pas en un seul point commun**. En d'autres termes, **elles ne présentent pas de point isobestique**. Ce phénomène est très important car il prouve sans ambiguïté que **la solution contient des espèces autres que les protéines native et complètement dépliée**, et qu'il existe nécessairement des états intermédiaires.

En effet, l'équation (.1.52) montre que, si la transition est à deux états, natif N et dénaturé D, on a pour toute valeur de la température :

$$\frac{I(Q_0, T)}{I(0, T)} = (f_N(T) + f_D(T))P_N(Q_0) = P_N(Q_0)$$

où Q_0 est le point de croisement des facteurs de forme $P_N(Q)$ et $P_D(Q)$, correspondant aux protéines natives et dénaturées. Les courbes normalisées $I(Q,T)/I(0,T)$, obtenues à différentes températures, se coupent donc toutes en Q_0 . Inversement, l'absence de point de croisement commun implique l'existence d'autres états. Il est important de souligner que cette absence de point isobestique n'a un sens que si l'on regarde les courbes normalisées $I(Q,T)/I(0,T)$, et non

les intensités diffusées $I(Q,T)$. En effet, si l'on reprend l'équation (1.50) pour une transition à deux états :

$$I(Q,T) = f_N(T)I_N(Q,T) + f_D(T)I_D(Q,T)$$

Au point de croisement Q_0 des intensités native et dénaturée, mesurées à des températures encadrant la transition, respectivement T_N et T_D , on a : $I_N(Q_0, T_N) = I_D(Q_0, T_D)$. Mais $I_N(Q_0, T)$ n'est pas égal à $I_N(Q_0, T_N)$, de même, $I_D(Q_0, T) \neq I_D(Q_0, T_D)$, à cause du terme de contraste qui varie avec la température. Il s'ensuit que les différentes courbes $I(Q,T)$ ne présentent pas forcément de point isobestique, même si la transition est à deux états. L'utilisation des courbes normalisées $I(Q,T)/I(0,T)$ est donc indispensable si l'on veut relier l'absence de point isobestique à l'existence d'états intermédiaires.

Il faut aussi souligner que le raisonnement précédent s'appuie sur l'hypothèse que le facteur de forme $P_i(Q)$ d'une espèce i ne varie pas avec la température. L'absence de point isobestique n'aura donc un sens que si l'on examine un domaine réduit en température, dans lequel on peut raisonnablement supposer que $P_i(Q)$ est constant. Nous sommes dans ce cas, puisque nous examinons le domaine en température [65°C - 75°C].

Pour finir, nous sommes conscients que l'on peut faire l'objection que les courbes présentées dans Fig4. 14 n'ont pas été corrigées des effets de concentration. Pour prouver que l'absence de point isodiffusif n'est pas une conséquence des interactions entre protéine, nous avons également porté en représentation de Kratky (Fig4. 15) les courbes extrapolées à concentration nulle, en utilisant l'équation (1.74), et en faisant la même hypothèse sur le coefficient du Viriel A_2 que dans la section 4.1.3.1. On observe que la correction est en fait très faible, et qu'il n'y a toujours pas de point isodiffusif. Dans la suite, nous travaillerons sur les courbes non corrigées, compte tenu du fait que la correction est très faible et qu'elle nécessite plusieurs approximations. Il nous a donc paru préférable de ne pas la faire.

Nous avons ensuite développé une analyse des spectres de diffusion, permettant de caractériser les états intermédiaires détectés en représentation de Kratky. C'est ce que nous exposons dans le paragraphe suivant.

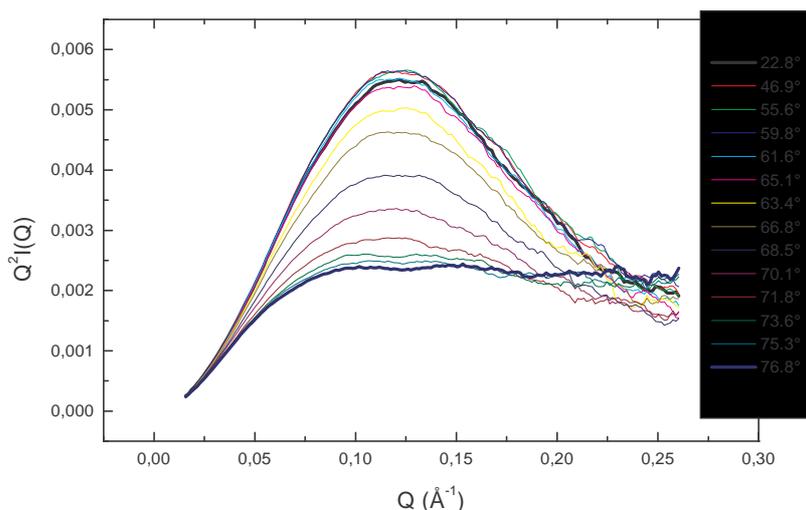


Fig4. 15 Représentation de Kratky des courbes extrapolées à concentration nulle.

4.1.4 Mise en évidence des états intermédiaires de repliement dans la transition en fonction de la température.

Pour tenter de déterminer le nombre minimal d'états intermédiaires de la NCS, qui interviennent dans le processus de dénaturation par chauffage, nous avons utilisé une méthode analytique, dite de décomposition en valeurs et vecteurs propres (Singular Value Decomposition, SVD, en anglais). C' est une méthode très générale qui permet de déterminer le nombre d' espèces indépendantes que l' on peut détecter par une technique expérimentale donnée (par exemple la diffusion aux petits angles, Provencher et al., 1983) au cours d' un processus de transition, et qui ne nécessite aucune connaissance *a priori* sur les espèces individuelles. Cette méthode a été utilisée plusieurs fois pour décrire des transitions structurales suivies par DXPA (Fetler et al., 1995), et en particulier dans le cas de processus de dénaturation (Chen et al, 1996, 1998; Segel et al., 1998, 1999, Koide et al., 1999)

4.1.4.1 Théorie de la méthode SVD

Le principe de base de cette méthode consiste à considérer l' ensemble des N courbes de diffusion expérimentales, qui comportent chacune M points de mesure, comme les éléments d' une matrice A de dimension $M \times N$. La colonne n^{e} de cette matrice représente la courbe de diffusion $I(Q, T_k)$ mesurée à la température T_k , et normalisée par l' intensité à l' origine $I(0, T_k)$. La matrice A peut être décomposée en un produit de trois matrices :

$$A(M \times N) = U(M \times N) w(N \times N) b(N \times N) \quad (4.3)$$

La matrice $U(M \times N)$ est constituée de N colonnes $u_j(Q)$ qui sont les vecteurs de base orthogonaux. La matrice $w(N \times N)$ est diagonale et contient les valeurs propres positives w_j de la décomposition, classés en ordre décroissant. Enfin la matrice $b(N \times N)$ contient les projections normalisées b_j^k de la courbe expérimentale $n^{\circ}k$ sur le vecteur de base u_j .

Dans notre cas, l'analyse SVD a été appliquée aux profils de Kratky, plutôt qu'aux courbes expérimentales $I(Q)$, dans le but de privilégier la région des plus grandes valeurs de Q , qui contient la plupart des informations structurales qui nous intéressent. Si l'on utilise l'équation 4.3, chaque profil de Kratky $n^{\circ}k$ peut se mettre sous la forme :

$$Q^2 I(Q, T_k) = \sum_{j=1}^N w_j b_j^k u_j(Q) \quad (4.4)$$

En pratique, les valeurs propres w_j de rang j élevé sont très faibles, et les projections b_j^k correspondantes sont distribuées aléatoirement autour de zéro en fonction de k . Seuls les L premiers termes de la somme ci-dessus vont contribuer significativement à la reconstruction des profils expérimentaux suivant l'équation (4.4). Les $N-L$ termes suivants constituent le bruit de fond. Le nombre L de termes contribuant à la reconstruction des courbes $I(Q, T_k)$ représente le nombre d'espèces indépendantes que l'on peut détecter au cours de l'expérience, en d'autres termes **le nombre minimal d'états thermodynamiques qui interviennent au cours du processus de transition**.

Le nombre L est sélectionné en examinant les valeurs des w_j et des b_j^k , le profil des vecteurs propres $u_j(Q)$, et en utilisant un χ^2 normalisé par les barres d'erreur expérimentales $\sigma(Q, T_k)$:

$$\chi_p^2 = \sum_{k=1}^M \sum_{j=1}^N \left[\frac{(I_{exp}(Q_j, T_k) - I_{calc,p}(Q_j, T_k))^2}{\sigma(Q_j, T_k)^2} \right] \quad (4.5)$$

où $Q^2 I_{calc,p}(Q_j, T_k)$ est la reconstruction de la courbe de Kratky $n^{\circ}k$ utilisant les p premières composantes de l'équation (4.4). L représente la valeur minimale de p pour laquelle $\chi_p^2 \leq 1$.

En résumé, la méthode SVD est une technique mathématique qui permet de déterminer, sans approximation et sans modèle, le nombre minimal de populations différentes apparaissant par exemple au cours d'une transition. Il faut souligner que les vecteurs propres $u_j(Q)$ extraits de l'analyse par SVD ne représentent pas les courbes de diffusion des L espèces indépendantes. L'obtention de ces courbes de diffusion constitue une deuxième étape de l'analyse, qui requiert une procédure d'ajustement et l'utilisation d'un modèle thermodynamique que nous allons maintenant décrire.

Nous nous plaçons dans le cas où trois espèces seulement sont présentes dans la solution pendant la dénaturation (ce qui est notre cas, comme nous le verrons dans la partie des résultats), c'est-à-dire $L=3$. Ces trois espèces correspondent aux trois états thermodynamiques de la protéine, natif (N), intermédiaire (I) et dénaturé (D), et sont caractérisées par les fonctions de diffusion normalisées $I_N(Q)$, $I_I(Q)$ et $I_D(Q)$ respectivement.

A chaque température T_k , le profil de diffusion mesuré expérimentalement, $I(Q, T_k)$, est la superposition des profils de diffusion de ces trois espèces, pondérés par leurs concentrations relatives :

$$I(Q, T_k) = f_N^k I_N(Q) + f_I^k I_I(Q) + f_D^k I_D(Q) \quad (4.6)$$

où $f_N^k + f_I^k + f_D^k = 1$.

On peut par ailleurs décomposer le profil de Kratky obtenu à chaque température suivant les vecteurs de base de SVD :

$$Q^2 I(Q, T_k) = w_1 b_1^k u_1(Q) + w_2 b_2^k u_2(Q) + w_3 b_3^k u_3(Q) \quad (4.7)$$

Le même type de décomposition peut être appliqué aux profils de Kratky correspondant aux trois espèces indépendantes N, I et D :

$$\begin{aligned} Q^2 I_N(Q) &= w_1 b_1^N u_1(Q) + w_2 b_2^N u_2(Q) + w_3 b_3^N u_3(Q) \\ Q^2 I_I(Q) &= w_1 b_1^I u_1(Q) + w_2 b_2^I u_2(Q) + w_3 b_3^I u_3(Q) \\ Q^2 I_D(Q) &= w_1 b_1^D u_1(Q) + w_2 b_2^D u_2(Q) + w_3 b_3^D u_3(Q) \end{aligned} \quad (4.8)$$

Puisque les vecteurs propres de SVD sont indépendants, la décomposition de toute fonction de diffusion sur les vecteurs de base est unique. Par conséquent, la combinaison des

équations (4.6) à (4.8) permet d'établir à chaque température une relation entre les projections b_j^k issues de l'analyse SVD, et les fractions de populations f_N^k , f_I^k et f_D^k :

$$b_j^k = f_N^k b_j^N + f_I^k b_j^I + f_D^k b_j^D \quad \text{pour } j=1,2,3 \quad (4.9)$$

Afin de satisfaire l'équation de sommation $f_N^k + f_I^k + f_D^k = 1$, les fractions de population des trois espèces sont réécrites sous la forme :

$$\begin{aligned} f_N^k &= \frac{1}{1 + e^{-\Delta G_1(T_k)/RT_k} + e^{-\Delta G_2(T_k)/RT_k}} \\ f_I^k &= \frac{e^{-\Delta G_1(T_k)/RT_k}}{1 + e^{-\Delta G_1(T_k)/RT_k} + e^{-\Delta G_2(T_k)/RT_k}} \\ f_D^k &= \frac{e^{-\Delta G_2(T_k)/RT_k}}{1 + e^{-\Delta G_1(T_k)/RT_k} + e^{-\Delta G_2(T_k)/RT_k}} \end{aligned} \quad (4.10)$$

où $\Delta G_1(T_k)$ et $\Delta G_2(T_k)$ représentent les changements d'énergie libre à la température T_k entre l'état natif et l'état intermédiaire, et entre l'état natif et l'état dénaturé, respectivement.

Il est alors possible d'obtenir les paramètres thermodynamique $\Delta G_1(T_k)$ et $\Delta G_2(T_k)$, et les projections b_j^N , b_j^I et b_j^D en minimisant le χ^2 suivant :

$$\chi^2 = \sum_{k=1}^N \sum_{i=1}^L \left[\frac{(b_{i,obs}^k - b_{i,calc}^k)^2}{(\sigma_i^k)^2} \right] \quad (4.11)$$

où les $b_{i,obs}^k$ et les $b_{i,calc}^k$ sont respectivement les coefficients de projection extraits de l'analyse par SVD, et calculés à partir de l'équation (4.9). Les variances σ_i^k sont déduites des variances expérimentales $\sigma(Q, T_k)$ en utilisant l'expression :

$$w_i^2 (\sigma_i^k)^2 = \sum_{j=1}^M (Q_j^2 \sigma(Q_j, T_k))^2 u_i(Q_j)^2 \quad (4.12)$$

On peut alors, à l'aide de l'expression (4.8), reconstruire les profils de Kratky correspondant aux trois états N, I et D.

En résumé, cette deuxième étape de l'analyse permet d'obtenir *l'évolution des fractions de population des différentes espèces mises en jeu au cours de la transition de dénaturation, ainsi que les courbes de diffusion associées à ces espèces.*

4.1.4.2 Résultats

Comme mentionné précédemment, nous avons utilisé pour faire l'analyse par SVD les représentations de Kratky des courbes de diffusion $I(Q, T_k)$ normalisées à l'intensité à l'origine $I(0, T_k)$. Les courbes de diffusion que nous avons prises en compte sont celles obtenues pour les températures comprises entre 61.6°C et 76.8°C. Elles sont au nombre de dix. La raison pour laquelle nous avons travaillé sur un domaine en température réduit est la même que celle exposée en section 4.1.3.3 dans la discussion sur l'absence de point isobestique : l'analyse par SVD n'aura un sens que si le facteur de forme (autrement dit l'intensité de diffusion normalisée) d'une espèce i donnée ne varie pas dans le domaine en température exploré. C'est pourquoi nous avons réduit ce domaine au voisinage immédiat de la transition.

Le premier travail a donc consisté à sélectionner le nombre minimal d'espèces indépendantes nécessaire pour décrire l'ensemble des courbes de diffusion, en calculant le χ^2 donné par l'équation (4.5), et en examinant l'allure des vecteurs de base $u_j(Q)$ et le comportement en fonction de T_k des coefficients $w_j b_j^k$ de ces vecteurs.

La figure 4.16 donne les représentations de Kratky des trois premiers vecteurs de base, la figure 4.17 représente l'évolution en température des coefficients $w_j b_j$ des vecteurs de base u_j pour j allant de 1 à 4.

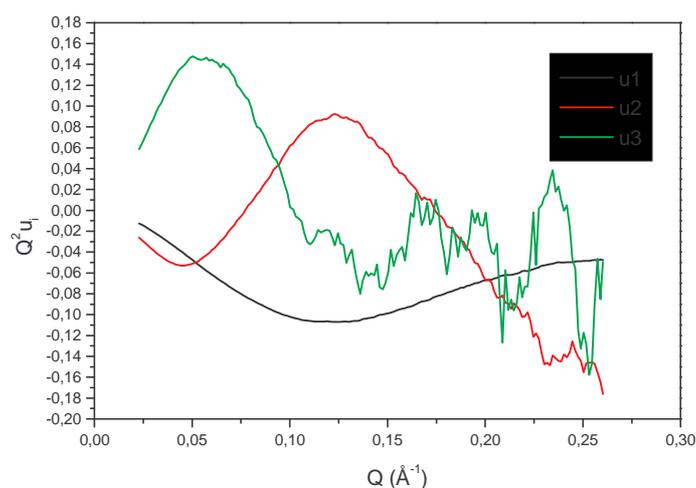


Fig4. 16 Représentation de Kratky des trois premiers vecteurs de base $u_j(Q)$.

Ces figures montrent immédiatement que *le nombre minimal d'espèces indépendantes nécessaire pour décrire la transition de dénaturation est trois*, puisque les trois premiers vecteurs propres u_j , ainsi que les coefficients $w_j b_j$ associés, adoptent des valeurs significatives. Par contre les valeurs $w_4 b_4$ correspondant au quatrième terme sont distribuées aléatoirement autour de zéro en fonction de la température T_k . De la même façon, le quatrième vecteur propre $u_4(Q)$ (non représenté sur la figure 4.16) fait partie du bruit de fond expérimental. On a vérifié également que χ_2^p devient égal à 1 pour $p=3$.

L'analyse par SVD démontre donc qu'*il existe au moins un état intermédiaire*, différent des états natif et complètement déplié, au cours de la transition de dénaturation par chauffage. Elle ne démontre pas qu'il en existe un seul. On peut seulement dire que, *à la précision de nos mesures près*, le quatrième vecteur propre, qui correspondrait à l'existence d'un second état intermédiaire détectable, ne contribue pas de façon significative à la reconstruction des courbes de diffusion. Il faudrait des données avec une meilleure statistique pour en mettre en évidence d'éventuelles espèces encore moins peuplées que l'état intermédiaire mis en évidence dans la présente analyse.

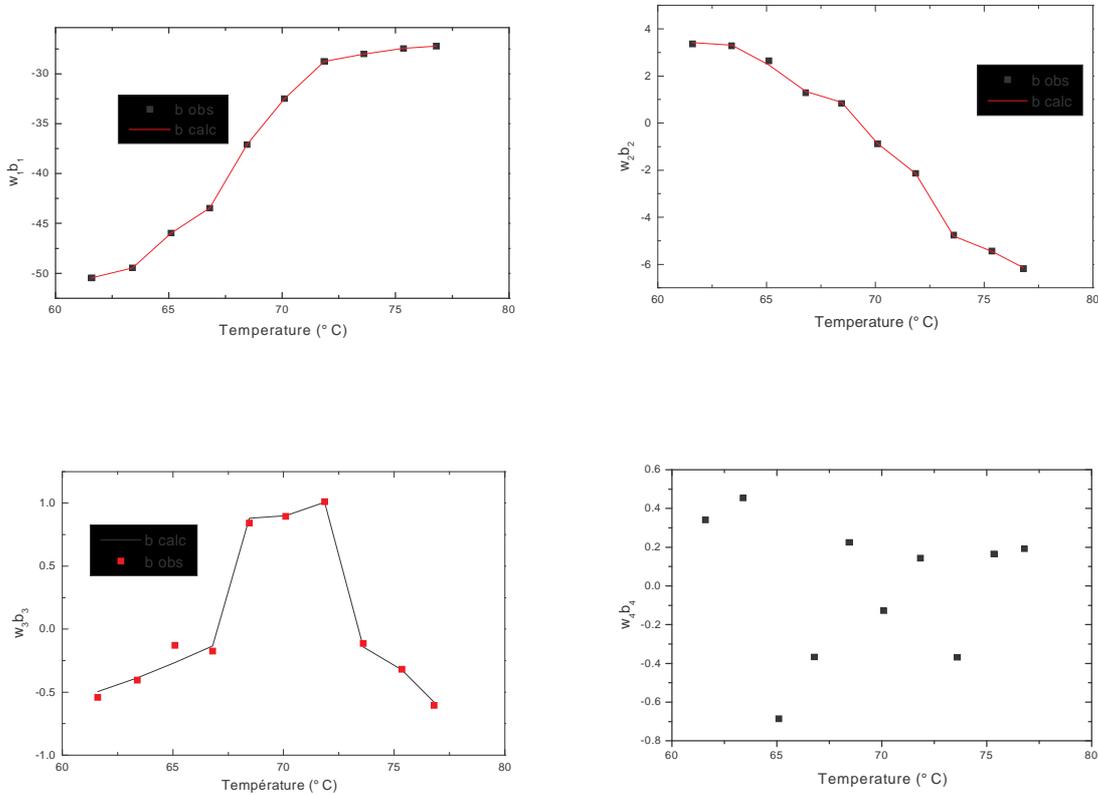


Fig4. 17 Evolution en fonction de la température des coefficients $w_j b_{j,obs}$ des quatre premiers vecteurs de base, extraits de l'analyse par SVD. Le trait solide représente le résultat $w_j b_{j,calc}$ des ajustements de ces coefficients à l'équation 4.7.

Dans une deuxième étape, nous avons ajusté à l'équation (4.10) les coefficients de projection $b_{i,obs}^k$ issus de l'analyse par SVD, ce qui nous a permis d'obtenir d'une part l'évolution avec la température des fractions de population des trois espèces N, I, et D, et d'autre part les valeurs des coefficients b_j^N , b_j^I et b_j^D ($j=1$ à 3). La détermination de ces coefficients nous a ensuite permis de reconstruire les courbes de diffusion $I_N(Q)$, $I_I(Q)$ et $I_D(Q)$ correspondant aux trois espèces, à partir de l'expression (4.9). Il a été nécessaire d'imposer deux contraintes pour réaliser l'ajustement, étant donné le grand nombre de variables : $f_N=1$ à $T=61.6^\circ\text{C}$, et $f_D=1$ à $T=76.8^\circ\text{C}$. La première contrainte est justifiée dans la

mesure où la courbe de diffusion expérimentale mesurée à 61.6°C et extrapolée à concentration nulle est parfaitement décrite à l'aide du programme CRY SOL en prenant les coordonnées cristallographiques de la protéine native. La deuxième contrainte est justifiée par le fait que la courbe de diffusion à 76.8°C est analogue à celle d'une chaîne de Kratky-Porod, et également par le fait que les courbes de diffusion ne présentent pas d'évolution entre 76.8 et 80°C, ce qui laisse supposer que la protéine a déjà atteint à 76.8°C son état complètement déplié.

La figure 4.18 montre la variation en température des fractions des populations native, intermédiaire et dénaturée. On observe que l'espèce intermédiaire atteint son maximum de population (de l'ordre de 40%) entre 68 et 72°C, c'est-à-dire au voisinage du point de demi-transition. Jusqu'à la température de 68°C, l'état dénaturé est peu peuplé, et la solution contient essentiellement des protéines natives et des protéines dans l'état intermédiaire. Par contre, à partir de 72°C, les espèces prédominantes sont dans l'état natif et dans l'état intermédiaire.

Les valeurs des variations d'énergie libre $\Delta G_1(T_k)$ et $\Delta G_2(T_k)$, associées suivant l'équation (4.12) aux valeurs des fractions f_N^k , f_I^k et f_D^k données dans la figure 4.18, présentent un comportement approximativement linéaire en fonction de la température. Si on fait l'hypothèse que l'on peut extrapoler ces valeurs jusqu'à 300K, on obtient pour ΔG_1 et ΔG_2 à 300K des valeurs de l'ordre de 7kcal/mol et 2kcal/mol respectivement. L'ordre de grandeur de ces paramètres est très satisfaisante pour une protéine globulaire de petite taille.

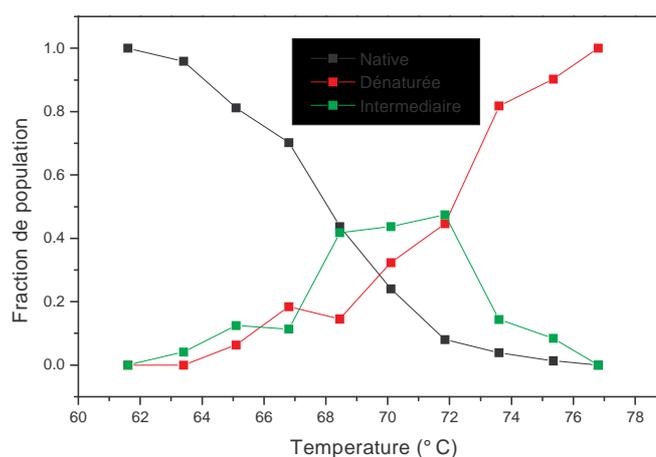


Fig4. 18 Fractions des populations native, intermédiaire et dénaturée en fonction de la température.

Les représentations de Kratky des spectres de diffusion reconstruits $I_N(Q)$, $I_I(Q)$ et $I_D(Q)$, correspondant aux trois espèces native, intermédiaire et dénaturée, sont présentées sur la figure 4.19. Les spectres $I_N(Q)$ et $I_D(Q)$ sont évidemment analogues à ceux mesurés expérimentalement pour les protéines native et dénaturée, respectivement, puisque nous avons imposé les contraintes $f_N=1$ à $T=61.6^\circ\text{C}$, et $f_D=1$ à $T=76.8^\circ\text{C}$. La représentation de Kratky de la courbe de diffusion reconstruite $I_I(Q)$ montre que l'espèce intermédiaire est relativement proche de l'état complètement dénaturé, c'est-à-dire certainement assez déstructurée, mais conserve encore certains caractères de globularité, comme l'atteste la présence du pic large centré vers 0.11\AA^{-1} . Le rayon de giration déterminé à partir de la courbe de diffusion reconstruite $I_I(Q)$, en utilisant la loi de Debye, est le suivant :

$$R_{gI} = 23.4 \pm 0.1 \text{ \AA}$$

à comparer à la valeur de 14 \AA pour la protéine native, et 26.2 \AA pour la protéine dénaturée.

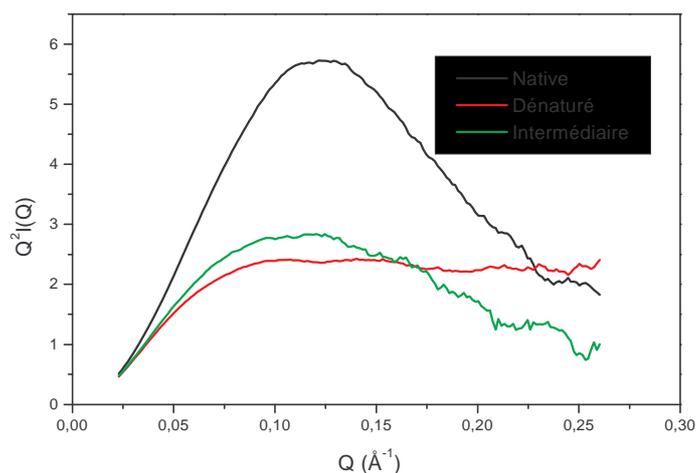


Fig4. 19 Représentations de Kratky des spectres de diffusion reconstruits $I_N(Q)$, $I_I(Q)$ et $I_D(Q)$, correspondant aux trois espèces native, intermédiaire et dénaturée.

En conclusion, l'analyse par SVD des courbes de diffusions aux petits angles de la NCS confirme la présence d'états intermédiaires dans le processus de dépliement, déjà subodorée par l'absence de point isodiffusif dans la représentation de Kratky des courbes de diffusion. L'espèce qui a été mise en évidence est très déstructurée, avec un rayon de giration élevé, mais avec également un degré de globularité non négligeable. Il est important toutefois de souligner que notre analyse n'est pas incompatible avec l'existence de plusieurs états

intermédiaires, voire une distribution d'états intermédiaires, tous relativement déstructurés. L'état intermédiaire que nous avons caractérisé au cours de notre analyse pourrait alors représenter une sorte de moyenne de cette distribution d'états.

4.2 Transition de dénaturation, à l'équilibre, en fonction du Chlorure de guanidinium

Pour caractériser la transition structurale de dénaturation en chlorure de guanidinium, nous avons comparé les informations données par la fluorescence, le dichroïsme circulaire et la diffusion des neutrons aux petits angles. Chacune de ces techniques a apporté des informations complémentaires sur la variation de la structure de la NCS au cours de la dénaturation. Rappelons encore une fois que la caractérisation a été faite soit en H₂O soit en D₂O.

4.2.1. Variation de la structure secondaire et tertiaire en fonction du chlorure de guanidinium

La Fig4. 20 montre les courbes de dénaturation de la NCS, obtenues soit en H₂O ou en D₂O, suivies par déplacement de la longueur d'onde maximale d'émission de fluorescence des tryptophanes, en fonction de la concentration de dénaturant. Elles sont bien représentées par une simple transition et caractérisent la perte progressive de la structure tertiaire locale lorsque la concentration de dénaturant augmente.

En H₂O la protéine reste dans un état natif tant que la concentration de dénaturant reste inférieure à 2M de GdnHCl. La transition, se situe entre 2 M et 4 M de dénaturant, avec un point de demi-transition à 2.6M. Par contre en D₂O la protéine est plus stable. La NCS maintient sa conformation native jusqu'à 3M de GdnHCl, la transition plus coopérative s'effectue entre 3 M et 4 M avec un point demi-transition à 3.5 M. Dans les deux solvants la protéine perd complètement sa structure tertiaire locale en 4M de GdnHCl.

La longueur d'onde d'émission maximale varie entre 343 nm et 350 nm entre l'état natif et l'état dénaturé de la protéine. La valeur élevée de la longueur d'onde de la protéine native est due à la forte exposition au solvant des deux tryptophanes. En effet, la structure cristallographique montre très clairement que les deux tryptophanes sont fortement exposés au solvant (fig 1.8). Dans cette configuration, les fluorophores peuvent explorer un grand nombre

de conformations, produisant un spectre d'émission très déstructuré et fortement déplacé vers le rouge. La valeur de la longueur d'onde acquise par la protéine complètement dénaturée correspond à la longueur d'onde caractéristique d'un tryptophane totalement exposé au solvant.

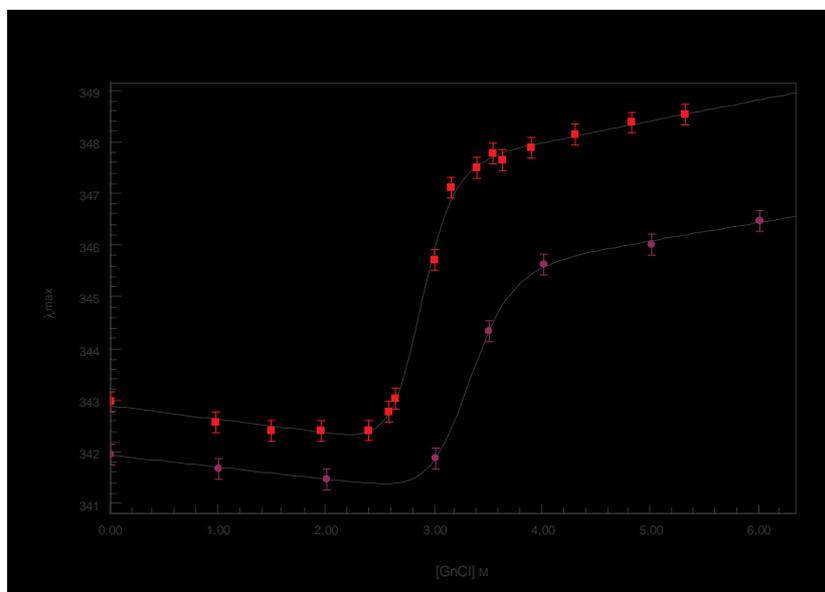


Fig4. 20 Courbes de dénaturation établies par fluorescence de tryptophanes

La variation du taux de structure secondaire de la NCS, est caractérisé par le signal dichroïque dans le domaine du lointain UV (Fig4. 21). En raison de l'impossibilité de suivre le signal dans cette région du spectre, à cause de la forte absorption de la lumière par le GdnHCl, la mesure de l'ellipticité a été faite à 223 nm, région où se situe un pic d'ellipticité, caractéristique des contraintes internes de la chaîne polypeptidique (voir paragraphe 3.1.2).

Comme pour les résultats de fluorescence, le dichroïsme circulaire indique également une plus grande stabilité de la protéine en eau lourde. Le signal dichroïque en H₂O, est constant jusqu'à 2M. Pour des concentrations supérieures, le signal dichroïque diminue, ce qui suggère le début de perte de structure secondaire. La perte de l'ellipticité est totale à la concentration de 3.5M.

En tampon deutéré, la protéine semble avoir une transition décalée de presque un molaire de dénaturant.

Dans les deux tampons, à la concentration de 5 M la protéine est dénaturée et une seule transition structurale est identifiée. Elles sont caractérisées par des points demi-transition de 2.8 M en H₂O et 3.4 M en D₂O.

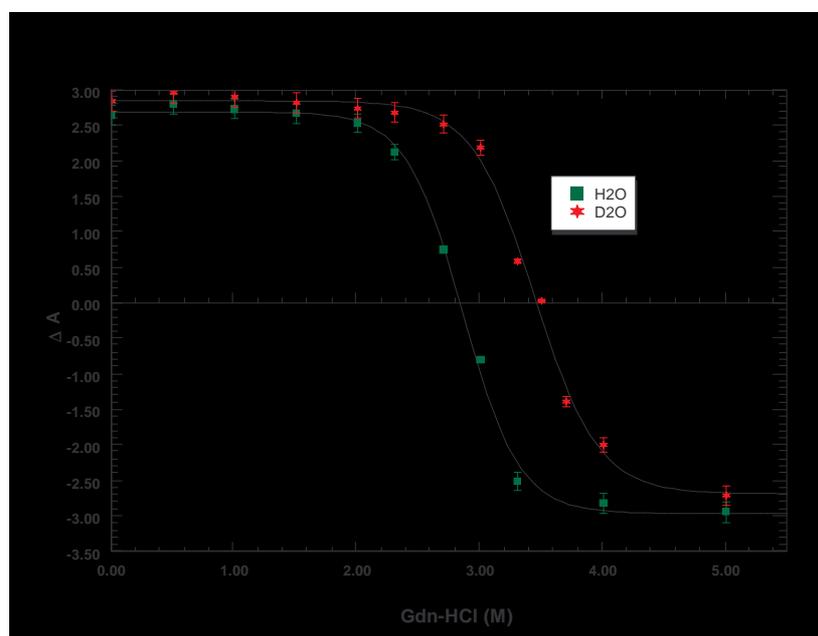


Fig4. 21 Courbes de transition établies par CD, en fonction du chlorure de guanidinium et du solvant.

4.2.2 Détermination du rayon de giration par diffusion aux petits angles des neutrons.

Les résultats précédents ont permis de caractériser les changements de structure secondaire et de structure tertiaire locale au cours de la transition. L'étape suivante a été d'acquérir des informations sur l'évolution de la forme globale et de la taille de la molécule, par diffusion de neutrons aux petits angles. Malheureusement, la durée de telles expériences est considérable et il n'a pas été possible de faire des acquisitions dans la région des grandes valeurs de Q pour les différentes concentrations en GdnHCl de la zone de transition. Nous disposons donc seulement de données à basse résolution, c'est-à-dire avec un vecteur de diffusion maximal de 0.1 \AA^{-1} . Dans ces conditions, nous avons pu obtenir la courbe de transition du rayon de giration, mais nous n'avons pas pu faire une analyse plus poussée des courbes de diffusion, comme nous l'avions fait en rayons X.

La Fig4. 22 représente les spectres de diffusion de la NCS, pour différentes valeurs de la concentration en GdnHCl. Les mesures ont été effectuées pour une concentration de

protéine de 5 mg/ml. Les résultats présentés ci-dessous n'ont pas été corrigés de l'effet des interactions, puisque nous ne disposons pas de données sur le comportement des coefficients A_2 et B_2 dans la région de transition. De toutes façons, comme nous l'avons déjà mentionné, ces effets ont peu d'influence sur l'allure des courbes de diffusion et les valeurs du rayon de giration quand on travaille à 5 mg/ml.

Tant que la concentration en GdnHCl reste inférieure à 3.3M, les spectres de diffusion ont la même allure que celui de la protéine native. Dans la plage en Q^2 comprise entre 0.001 \AA^{-2} et 0.01 \AA^{-2} , $\text{Ln}I(Q)$ varie linéairement en fonction de Q^2 , et le rayon de giration peut être déterminé à partir de l'approximation de Guinier. On constate que R_g varie peu en fonction de la concentration en GdmHCl jusqu'à 3.3M, la protéine gardant une forme compacte. Au-delà, les courbes de diffusion changent d'allure et l'utilisation de la loi de Debye est requise pour déterminer R_g .

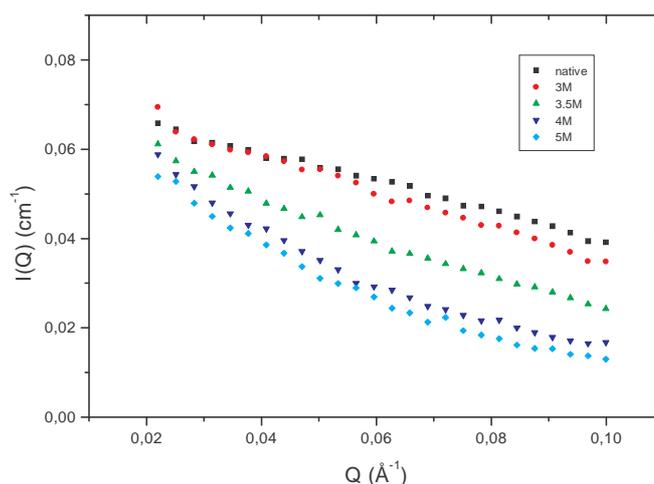


Fig4. 22 Spectres de diffusion de neutrons de la NCS dénaturée par le Chlorure de guanidinium.

La figure 4.23 présente la courbe de transition du carré du rayon de giration. Il s'agit du rayon de giration apparent, non corrigé des effets de concentration. Dans l'intervalle de concentration de GdnHCl [0M , 6M], le rayon de giration varie entre 12.5 \AA et 30.6 \AA . La transition se déroule avec un point de demi-transition à 3.6M, très proche des valeurs des points de demi-transition détectés par fluorescence et par dichroïsme circulaire en tampon deutérié.

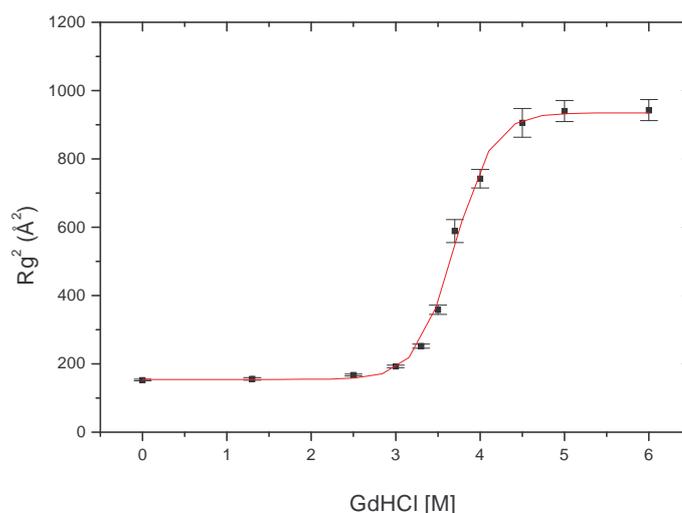


Fig4. 23 Courbe de transition, en fonction de la concentration en GdmHCl, du carré du rayon de giration de la NCS, déterminé pour une concentration de protéine de 5mg/ml.

Les trois transitions de denaturation que nous venons de décrire ont la caractéristique commune d'être symétriques, de présenter seulement une transition structurale et donc de pouvoir être décrites, en première approximation, comme de simples transitions à deux états. La stabilité conformationnelle de la NCS est caractérisée par les points de demi-transition et peut être estimée à partir de ces transitions. Les paramètres thermodynamiques ont été calculés pour chacune de ces courbes, en utilisant l'approximation d'un processus à deux états et d'après le modèle de Pace décrit dans le chapitre II.

Dans cette approximation, l'énergie de Gibbs et la valeur de la constante de proportionnalité m , pour la transition détectée par *fluorescence*, sont évaluées respectivement à: 9.9 ± 1 kcal/mol and 2.9 ± 0.5 kcal/mol en D_2O et à 11.7 ± 1 kcal/mol and 4.1 ± 0.4 kcal/mol en H_2O .

Par la courbe de *dichroïsme* nous avons estimé un ΔG de 7.2 ± 0.7 kcal/mol et m de 2.5 ± 0.2 en H_2O , et de 7.8 ± 0.7 kcal/mol et 2.4 ± 0.2 kcal/mol en D_2O , respectivement.

L'analyse de la courbe de transition du carré du rayon de giration a donné des valeurs de l'énergie libre égale à 9.06 ± 1.2 kcal/mol et de m égale à 2.4 ± 0.2 kcal/mol.

Comme nous l'avons déjà vu pour la dénaturation thermique, la différence des paramètres thermodynamiques calculés à partir des courbes de transition de dénaturation établies par diverses techniques, nous suggère l'hypothèse que le modèle à deux états n'est pas suffisant pour décrire complètement la transition.

4.2.3 Mise en évidence des états intermédiaires de repliement dans la transition en fonction du chlorure de guanidinium.

Les données présentées dans le paragraphe précédent suggèrent fortement la présence d'intermédiaires dans le processus de dénaturation de la NCS. Afin de vérifier cette hypothèse, nous avons abordé le problème à l'aide d'une sonde fluorescente hydrophobe, l'ANS. Cette sonde nous permet de suivre l'exposition au solvant des résidus apolaires, situées au cœur de la protéine, au cours de la transition de dénaturation, et d'étudier les premières étapes du processus.

Le 8-anilo-1-naftalene-sulfonate (ANS) a la puissante propriété d'avoir un rendement quantique très bas dans l'eau et presque égal à 1 dans un environnement apolaire (Robinson, 1978), et de ne pas former de liaison covalente avec la protéine. L'ANS forme des liaisons électrostatiques avec les protéines, avec une enthalpie de liaison comprise entre -3kcal/mol et -6kcal/mol (Matulis 1998). La liaison de l'ANS dépend principalement de la charge de la macromolécule, sa fluorescence dépendant d'un certain nombre de facteurs structuraux comme par exemple le voisinage des molécules d'eau (Kirk, 1996)

Comme cela est attendu, la Fig4. 24 montre que l'ANS interagit très faiblement avec l'état natif et l'état complètement dénaturé, alors qu'une variation est clairement présente entre 1M et 3M avec un maximum à 2M. Ceci reflète le changement d'environnement de la molécule. En effet, quand la protéine est native, la sonde fluorescente est dans un environnement polaire et son rendement quantique est très faible. Quand la protéine est suffisamment dénaturée l'ANS peut entrer dans la matrice protéique et venir en contact avec les résidus hydrophobes. La molécule étant plus petite que la protéine (MW 299D) elle peut pénétrer dans la protéine si celle-ci présente une structure moins compacte que l'état natif. Par contre, à hautes concentrations de dénaturant, où, comme nous l'avons vu dans la première section, les structures secondaires et tertiaires sont absentes et les résidus hydrophobes

exposés au solvant, l'ANS reste dans un environnement polaire et son intensité de fluorescence diminue.

Il est clair que, dans le cas d'un d'équilibre à deux états entre l'état natif et l'état dénaturé, où l'ANS serait toujours exposé au solvant, l'intensité de fluorescence ne devrait avoir aucun changement en fonction du paramètre de dénaturation. Par conséquent, la variation de fluorescence observée est significative typique d'un chemin de dénaturation plus complexe.

Il est important d'observer que le maximum de fluorescence de l'ANS est obtenu pour des concentrations de dénaturant où les signaux de fluorescence, dichroïsme circulaire et rayon de giration indiquent une conformation encore native. Ces résultats semblent donc confirmer la présence d'un état intermédiaire compact. Il est nécessaire de souligner qu'aux longueurs d'onde de mesure, le signal de fluorescence de la protéine est nul est que seul celui de l'ANS est visible.

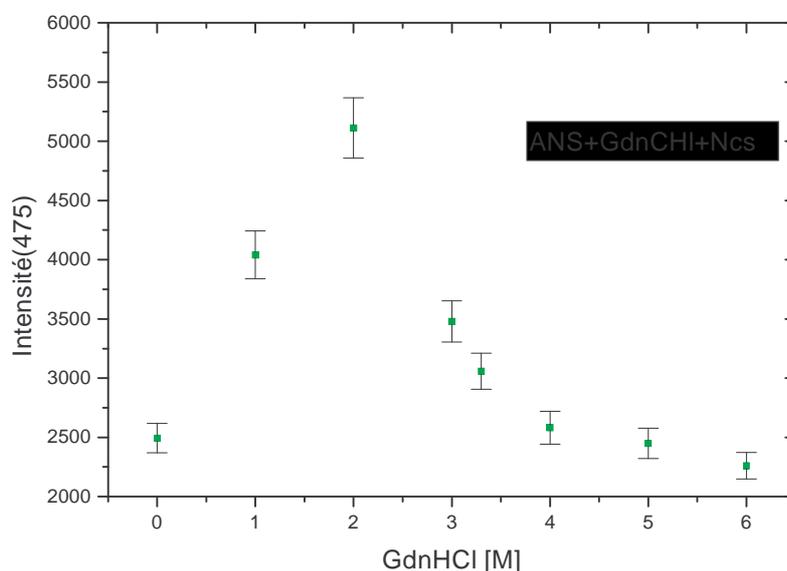


Fig4. 24 Variation de l'intensité de fluorescence de la ANS à 275 nm

4.2.4 Etude thermodynamique

Au vu des résultats précédents, nous avons entrepris une étude thermodynamique, afin de mieux caractériser cet intermédiaire et pour essayer de déterminer si nous sommes en présence d'une seule espèce stable où s'il s'agit d'un ensemble continu d'états dont le degré de compacité varie en fonction de la concentration de GdnHCl. L'idée principale a donc été de faire des mesures de microcalorimétrie différentielle en présence de concentration de dénaturant où nous avons observé la fluorescence de l'ANS. Nous avons ainsi suivi la dénaturation de la NCS en fonction de la température dans un intervalle de 25°C-80°C, en faisant varier son état initial en ajoutant 1M, 2M et 3M de GdnHCl en eau légère.

La Fig4.25 montre que, comme cela était attendu, la stabilité de la protéine diminue en présence du dénaturant. Pour les deux différents solvants, les amplitudes et les températures de demie transition diminuent lorsque la concentration du dénaturant augmente. Toutefois l'une des caractéristiques importantes de ces transitions est que leur profil présente une allure asymétrique qui devient plus marquée lorsque la concentration de dénaturant augmente. Ce phénomène n'est pas attribuable à la formation d'agrégats puisque les courbes de transition sont parfaitement réversibles et reproductibles pour chaque concentration de dénaturant utilisé.

Ces courbes de transition n'ont pu être analysées qu'au moyen de deux transitions *non à deux états*. Dans certains cas, et plus précisément quand la concentration de dénaturant est plus élevée, les courbes de transition nécessitent d'être considérées comme une convolution de plusieurs transitions, chacune d'entre elles étant une transition *non à deux-états*. En général, en présence d'une population d'intermédiaires peu peuplée, la transition que l'on mesure est le résultat d'une convolution des plusieurs transitions (Novokhanty, 1984) comme c'est le cas des échantillons en 2 M et 3 M Gdn-HCl.

Nous avons constaté que la convolution de deux transitions est suffisante pour un bon ajustement des courbes de fusion, à condition que chacune d'elle soit représentée par une convolution d'un nombre fini de transitions. L'impossibilité d'ajuster les courbes de DSC avec deux transitions à deux-états à la concentration de 1M suggère l'exclusion d'une situation où seules trois espèces seraient présentes au cours du dépliement.

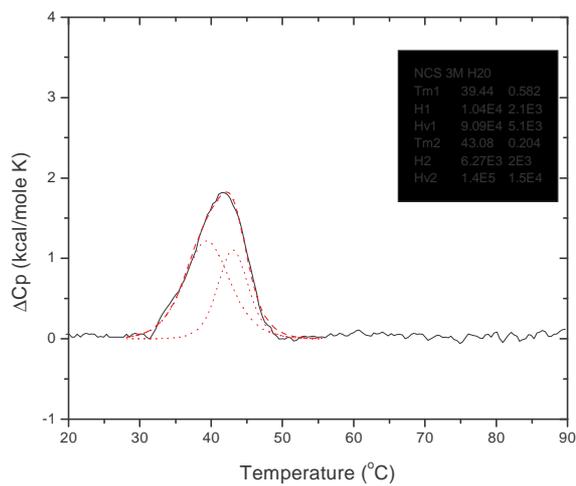
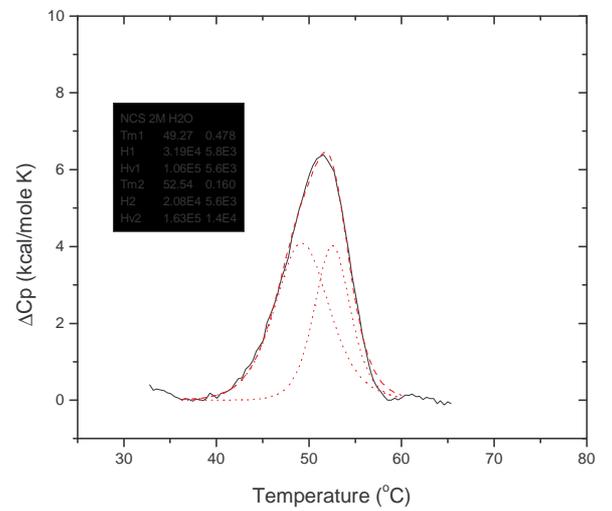
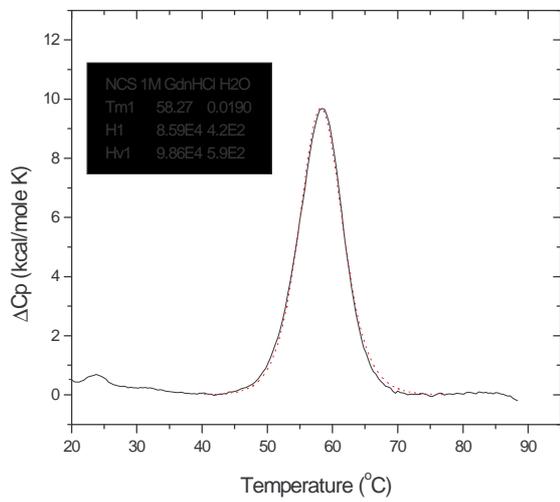


Fig4. 25

*Ajustement des données des courbes DSC de la NCS α
1M, 2 M, 3 M, de chlorure de guanidinium en H2O*

Les données expérimentales pourraient alors s'interpréter par la présence en solutions soit d'une seule espèce qui change en continu avec la température (ce qui pourrait correspondre à une transition du second ordre) soit de plusieurs espèces thermodynamiquement stables.

Dans tous les cas, ces résultats de microcalorimétrie confirment les résultats obtenus par fluorescence de l'ANS qui suggèrent l'existence d'intermédiaires dans la région *native-like* des courbes de transitions obtenues par d'autres techniques (rayon de giration, signal optique).

Pour avoir une complémentarité des données, il nous a semblé intéressant d'effectuer ces mesures en D₂O, en faisant l'hypothèse que le décalage observé entre eau légère et eau lourde est également valable pour les expériences d'ANS. C'est pourquoi nous avons décidé d'étudier deux échantillons l'un à 2M l'autre à 3M en eau lourde (Fig4. 26) afin de détecter les intermédiaires en D₂O.

Comme pour les expériences en H₂O, il a été important de contrôler la réversibilité des balayages pour vérifier qu'aucune formation d'agrégats n'influçait le thermogramme. Deux mesures consécutives à 2M de D₂O, à 1 K/min ont montré une réversibilité de 100%, montrant que la présence du chlorure de guanidinium prévient la formation d'oligomères.

Enfin, l'analyse des transitions de dénaturation, observées (Fig4. 26), suggère que les intermédiaires en eau lourde possèdent des propriétés équivalentes à celles des intermédiaires détectés en eau légère.



Fig4. 26 Ajustement des données des courbes DSC de la NCS en 2 M, 3 M, de chlorure de guanidinium en eau lourde.

Tous les paramètres thermodynamiques que nous avons obtenus au cours de cette analyse de calorimétrie, soit en H₂O soit en D₂O, sont listés en table 4.1. Les valeurs ont été estimées en considérant deux transitions successives pour la solution de NCS en 2 M et 3 M de chlorure de guanidinium et seulement une pour les autres.

Tableau 4.2

Echantillon	Tm ₁ (°C)	Hc ₁ (kcal/mol)	Hv ₁ (kcal/mol)	Tm ₂ (°C)	Hc ₂ (kcal/mol)	Hv ₂ (kcal/mol)
0 M H ₂ O	68.6	97.3	115			
1 M H ₂ O	58.27	85.9	98.6			
2 M H ₂ O	49.27	31.9	76	54.5	36.8	120
3 M H ₂ O	39.44	10	90	43	6.2	140
0 M D ₂ O	71.83	97.7	119			
2 M D ₂ O	49.19	30	102	52.6	21.7	140
3 M D ₂ O	39.83	15.5	84.6	44.08	15.5	125

Nous pouvons observer une diminution continue du Tm et de l'enthalpie lorsque la concentration de GdnHCl augmente. Les échantillons à 0 M sont ceux présentés dans le chapitre III.

Commentaire.

Comme le montrent les résultats des courbes de diffusion aux rayons X, un intermédiaire est aussi présent pendant le dépliement en fonction de la température. Nous nous sommes alors posés la question si pendant la caractérisation en DSC aux différentes concentrations de GdnHCl, l'espèce mise en évidence par SVD était encore présente et en quelle mesure elle pouvait influencer la caractérisation de l'intermédiaire compact. L'espèce intermédiaire qui intervient pendant le dépliement en température est plutôt déstructurée. Il est probable qu'en présence des concentrations de chlorure de guanidinium que nous avons utilisées, cette espèce n'arrive pas à se mettre en place. Dans ce cas, les effets observés seraient strictement reliés à l'intermédiaire que l'on veut caractériser.

Les mesures de calorimétrie reflètent une série très complexe de phénomènes qui incluent non seulement les changements de conformations mais aussi des effets d'ionisation de la protéine et du solvant et des effets de solvation de la protéine par le dénaturant (formation de liaisons hydrogènes sur des sites préférentiels). Ce dernier phénomène peut contribuer de façon non négligeable sur l'économie de l'enthalpie totale associée à la transition (Privalov 1979). En fait, il a été estimé que cette enthalpie de liaison pour le cytochrome c, la ribonuclease et le lysozyme en 6M de GdnHCl à 25°C, est de l'ordre de -2 kcal/mol par site. Pour le lysozyme, il est reconnu qu'à 6M il existe environ 65 sites de liaison, ce qui donne une contribution de la chaleur moyenne de solvation d'environ -130 kcal/mol, valeur loin d'être négligeable. On peut prendre en compte cette enthalpie d'interaction, qui peut être calculée connaissant le nombre des sites de liaison dans la protéine native et dans la protéine dénaturée (Privalov 1979). Mais, plus simplement, dans le cas d'un système à deux états, si l'on considère l'enthalpie mesurée comme la somme de celle de dénaturation et de celle d'interactions soluté-solvant, la différence entre la valeur à zéro molaire et à une concentration finie de GdnHCl nous donne une estimation de cette énergie de liaison.

Dans notre cas il est compliqué d'effectuer cette estimation, parce que nous ne connaissons pas la différence de ΔH qui existe entre l'état natif et 'intermédiaire'. La variation d'enthalpie entre 0 M et une quantité finie de concentration de dénaturant ne nous donnera pas l'énergie de liaison. Il est très important de comprendre que plus le degré de dénaturation est élevé plus les sites de liaison sont disponibles et plus l'enthalpie calorimétrique est sous-

estimée. Cette sous-estimation de l'enthalpie nous donne une sur-estimation de l'enthalpie de van't Hoff. Pour cette raison, il nous a semblé illusoire de vouloir chercher de comparer, les différentes enthalpies afin d'identifier la présence des états intermédiaire ou des oligomères.

Mais si les effets de solvation affectent la variation d'enthalpie et la diminution du T_m , ils n'influencent pas la forme de la courbe de la chaleur spécifique. L'asymétrie mise en évidence est donc sûrement dominée par la présence des états intermédiaires de repliement. Quand la population des états intermédiaires tend vers zéro, la forme de la courbe de chaleur spécifique se rapproche de celle d'une transition à deux états, comme c'est le cas pour les échantillons à 0 M et à 1 M de dénaturant. Une variation dans la stabilité des différents états entraîne dans des changements significatifs dans le profil des courbes de fusions.

A ce sujet, il existe des exemple qui montrent que si la transition est à deux états, les courbes de la chaleur spécifique sont symétriques. Nous pouvons citer un travail fait sur le repliement d'une petite protéine de 85 résidus, la HPr (histidine-containing phosphocarrier protein), à l'aide de la DSC, CD, ANS, et NMR (Van Nuland, 1998). Dans cet article, les auteurs étudient le processus de dénaturation de la protéine par le GdnHCl et par la température, en mettant en évidence que le dépliement peut être décrit dans les deux cas comme un processus à deux états. Ils montrent que les courbes de DSC en fonction du chlorure de guanidinium, pour des concentrations comprises entre 0 M et 1.5 M, présentent une diminution de la température de demi-transition et de l'enthalpie calorimétrique mais restent symétrique et analysables avec un modèle à deux états.

4.3 Conclusions.

Interaction protéine-solvant

Il nous semble important de revenir sur les résultats que nous avons présentés ainsi que sur l'ensemble des informations obtenues dans le chapitre précédent. En particulier, l'effet de la substitution de l'eau légère par l'eau lourde mérite quelques discussions.

En général, pour chaque méthode de dénaturation, la transition obtenue en D₂O est toujours caractérisée par des paramètres de dénaturation plus importants que ceux obtenus en H₂O. Ceci se traduit par une coopérativité plus grande et un décalage du C_m et du T_m, tels que la stabilité de la protéine est plus haute dans le solvant deutéré ?

Deux questions se posent alors : d'une part quelle est la cause exacte de cet effet et, d'autre part, est-il dépendant de la protéine étudiée ?

De façon générale, il semble que cet effet H₂O/D₂O ne soit pas protéine dépendant ; des travaux effectués au laboratoire sur la PGK ont conduit à des résultats similaires (Receveur, 1997). D'autres données de la littérature vont aussi dans le même sens (Bonneté, 1994 ; Kuhlman 1998, Goto 1993).

Il faut donc rechercher la cause de ce phénomène dans les interactions protéine solvant et dans les interactions à l'intérieur de la protéine, indépendantes de la nature de la structure de la protéine. Bien entendu, la liaison hydrogène est le premier type d'interaction qui pourrait rendre compte de cet effet : la liaison O-D...O est plus forte que la liaison O-H..O. Par ailleurs il est bien connu que les interactions hydrophobes sont aussi modifiées par la substitution H/D (Bonneté 1994). Il n'est pas possible de dire *a priori* si les effets observés proviennent d'une différence de la force de la liaison hydrogène quand l'hydrogène est substitué par un deutérium, ou d'une différence dans les interactions hydrophobes lorsque les groupes non polaires sont placés en D₂O. La complexité de l'effet H/D est illustrée par un nombre important de données présentes dans la littérature qui donnent lieu à des interprétations différentes, si ce n'est contradictoires.

Des informations sur le rôle relatif des liaisons hydrogènes protéine-solvant et à l'intérieur de la protéine ont été obtenues en comparant les comportements d'une protéine hydrogénée ou d'une protéine deutérée dans un solvant D₂O. Il apparaît que la stabilité et la cinétique de certaines protéines ne sont pas modifiées si le squelette devient deutéré (Litzhaki, 1996 ; Parker 1997)

Kresheck et al. (1965) ont par contre étudié les effets d'une substitution D-H sur l'interaction hydrophobe, en étudiant les paramètres thermodynamiques du transfert de l'eau légère dans l'eau lourde, de composés non polaires ou partiellement polaires. Les expériences ont montré que les interactions hydrophobes associés aux hydrocarbures (propane et butane) sont plus stables en H₂O qu'en D₂O et que, par contre, les forces hydrophobes associées aux chaînes latérales de certains acides aminés (alanine, phénylalanine) sont plus fortes en D₂O. Les auteurs en déduisent que, pour les protéines, les interactions hydrophobes qui mettent en jeu les chaînes latérales hydrogénées non polaires sont plus fortes en D₂O qu'en H₂O. Ce sont

probablement ces interactions qui, à 25°C, sont responsables du fait que la structure native est plus stable en D₂O qu'en H₂O.

Les facteurs qui déterminent la variation de la stabilité avec la température ont été discutés par Scheraga et al (1962). Ces auteurs ont montré que puisque l'énergie de transfert des chaînes latérales des acides aminés non polaires, de H₂O dans D₂O, devient plus grande en augmentant la température, alors les interactions hydrophobes sont plus fortes. Ceci se traduit par une plus grande stabilité de la protéine dans D₂O vis-à-vis d'une modification thermique.

La transition de dénaturation

Dans un premier temps nous avons analysé les courbes de transitions de la protéine, à l'équilibre thermodynamique, obtenues par augmentation de la température et ajout de chlorure de guanidinium. Cette étude nous a permis non seulement de caractériser la région intermédiaire de la transition de dénaturation, mais aussi de déceler la présence d'espèces intermédiaires de repliement.

Une caractéristique commune à toutes les courbes de transitions que nous avons obtenues par fluorescence, dichroïsme circulaire DNPA et DXPA, est qu'elles apparaissent symétriques avec une variation sigmoïdale du signal, interprétable en première approximation par une transition à deux états. Toutefois, si tout semble être consistant avec l'hypothèse d'un processus à deux états, c'est-à-dire par un simple équilibre tout ou long du processus entre l'espèce native et celle dénaturée, il ne faut pas oublier que ces signaux sont des signaux spectroscopiques qui traduisent une *moyenne* des populations présentes. Il est donc possible que soit ces courbes de transition correspondent une moyenne entre plusieurs transitions. Pour tenter de résoudre ce problème, il est intéressant de comparer les résultats obtenus par les différentes techniques. En effet quand le degré de dépliement est mesuré par plusieurs techniques, chacune sensible à une caractéristique particulière de la protéine, il peut devenir plus facile de comprendre quels sont les phénomènes qui caractérisent le dépliement.

La *transition en chlorure de guanidinium*, montre que les points de demi-transition semblent équivalents pour tous les signaux suivis. Plus clairement, en D₂O la transition a un C_m moyen de 3.6 M et en H₂O de 2.6 M. Cette correspondance pourrait signifier que la NCS commence à perdre sa structure secondaire, tertiaire et à changer de forme et de taille pour

une même concentration de chlorure de guanidinium. Cette hypothèse est fort réaliste, surtout si on raisonne en terme de transition à deux états où la valeur de C_m est la concentration de dénaturant à laquelle les protéines natives et dénaturées coexistent en même proportion.

Cette situation s'observe également en *température* où les courbes établies en fluorescence et diffusion centrale présentent exactement la même allure et les mêmes températures de demi-transition. Par contre, les données obtenues par CD suggèrent que la protéine perde sa structure secondaire avant la structure tertiaire. Cette hypothèse doit être modulée par le fait que la variation d'ellipticité mesurée à 223 nm est sensible essentiellement aux contraintes internes de la protéine et ne donne par des renseignements sur la variation de la structure secondaire principale, c'est-à-dire des feuillets β . Le même type de remarque peut être fait au sujet de la fluorescence : il faut rappeler que la NCS comporte deux tryptophanes très exposés au solvant, ce qui limite, d'une certaine façon, les informations locales sur la structure tertiaire que nous pourrions obtenir à partir du signal de fluorescence.

Enfin, il est évident que si l'équilibre de dénaturation de la protéine est vraiment un équilibre à deux états, les paramètres thermodynamiques, comme par exemple la variation de l'énergie libre de Gibbs, la variation d'enthalpie..., qui sont extrapolés des courbes de transitions, doivent être indépendants de la sonde utilisée et donc être identiques aux barres d'erreur expérimentales près. Or, les paramètres thermodynamiques estimés expérimentalement ne sont pas équivalents entre eux, ni dans le cas de la dénaturation thermique ni dans le cas de la dénaturation chimique. De plus, pour la protéine dénaturée thermiquement, l'absence d'un point isodiffusif dans les courbes de diffusion aux petits angles suggère fortement que la transition de dénaturation est plus complexe qu'un processus à deux états. Des études en DSC nous ont aussi montré que les courbes de fusion ne peuvent être ajustées correctement que si l'on considère au moins un troisième état dans la transition.

Les intermédiaires

L'ensemble des résultats ci-dessus nous a conduit à faire l'hypothèse qu'il existe des états intermédiaires à l'équilibre et nous avons choisi d'étudier plus précisément ces états en exploitant de façon extensive les courbes de diffusion en rayons X et en utilisant une sonde hydrophobe, l'ANS.

L'ensemble des informations structurales apportées par ces différents techniques nous a permis d'établir, en première approximation, que :

- un intermédiaire a été détecté dans les premières étapes du *processus de dépliement en fonction du chlorure de guanidinium*. Cet intermédiaire possède l'ensemble des caractéristiques structurales de la protéine native. Il serait compact, la structure secondaire restant intacte, seule la structure tertiaire étant faiblement modifiée, permettant l'intrusion de l'ANS à l'intérieur de la matrice protéique.
- pour le *processus de dénaturation thermique*, un intermédiaire a été détecté avec un maximum de population à environ 70°C, c'est-à-dire au voisinage du point de demi-transition. Celui-ci serait très déstructuré, avec un rayon de giration plus proche de l'état complètement dénaturé que celui de l'état natif. Toutefois il garde encore un certain degré de globularité.

L'état intermédiaire au cours du dépliement chimique

Les résultats de microcalorimétrie suggèrent que nous ne sommes pas en présence d'un intermédiaire homogène, et que la transition X – D est composite. En particulier, l'absence d'un deuxième pic de transition et la possibilité de définir de façon satisfaisante la transition principale comme une convolution de deux transitions *non-deux-états*, excluent *a priori* que l'intermédiaire soit une espèce unique et stable, style *molten globule* ou état partiellement dénaturé, comme celle décrit pour CheY (Filimonov, 1993), et l' α -lactalbumine (Kataoka).

Ceci s'apparente plutôt à une distribution de plusieurs états dénaturés, un comportement décrit par le modèle proposé par Alonso et Dill (1991) et qui représente une alternative au modèle de dénaturation à trois états. Le modèle de dépliement peut être pensé comme un processus à plusieurs états où il existe une seule barrière d'énergie pour aller de l'état natif à l'ensemble des états dénaturés qui eux, varient de façon continue d'une conformation compacte avec un cœur hydrophobe à une conformation complètement dénaturée en fonction de la concentration du dénaturant.

Ce modèle rend compte de la difficulté à déterminer un intermédiaire unique et de l'absence d'une population d'intermédiaires bien définie ainsi que de la variation progressive

et parallèle des différents signaux spectroscopiques. **La notion d'intermédiaire correspond alors à une distribution continue d'états dénaturés.** Dans ce schéma la distribution aux basses concentrations de dénaturant s'effectue autour d'états compacts comportant des structures résiduelles significatives. Aux plus fortes concentrations, la distribution se déplace vers des états moins structurés jusqu'à atteindre l'état complètement dénaturé.

Un tel schéma où il n'existe pas d'intermédiaire directement défini au niveau structural va à l'encontre de ce qui a été décrit pour d'autres protéines comme par exemple la PGK (Receveur 1997) ou pour l'OspA (Outer surface protein A) (Koide, 1999). En effet, par exemple, les informations structurales apportées par la diffusion des neutrons et par le dichroïsme circulaire pour l'étude de la PGK, protéine à deux domaines, ont conduit à l'hypothèse de la formation d'un intermédiaire à l'équilibre, compact et partiellement déstructuré. La déstructuration intervient à la surface de la protéine alors qu'un cœur hydrophobe, formé de feuillets β , reste intact. Le dépliement est décrit en deux étapes, de la surface vers le cœur. Il s'agit d'un effondrement hydrophobe partiellement déstructuré.

En revanche ce schéma s'accorde parfaitement à ce que l'on sait de la structure de la NCS. Le corps de la NCS est essentiellement composé du tonneau β et des boucles, il n'existe pas de structure qui puisse donner lieu à un état partiellement déplié dans les premières étapes de dénaturation pour former des espèces stables à l'équilibre. La protéine NCS présente une cavité hydrophobe formée par la couche interne du tonneau β et de deux boucles externes. Ce cœur restera intact, jusqu'à des concentrations de dénaturant où stabiliser les interactions hydrophobes deviendront trop faibles ; quand la concentration de dénaturant atteint une certaine limite, le scénario de la protéine qui se dénature peut correspondre au tonneau qui commence à perdre sa conformation, à se déformer jusqu'à se déstructurer complètement. Ainsi, la protéine pourrait changer continûment sa structure jusqu'à ce que le cœur devienne instable et « éclate » pour suivre encore une fois une variation continue pour atteindre enfin la conformation à volume exclu.

L'état intermédiaire au cours du dépliement thermique.

L'intermédiaire qui a été mis en évidence par l'analyse SVD, au cours du dépliement thermique, se présente très déstructuré, avec un rayon de giration élevé, mais avec également un degré de globularité non complètement négligeable. Il faut rappeler toutefois que le type

d'analyse que nous avons fait, n'est pas incompatible avec l'existence d'une distribution d'états intermédiaires, comme celle utilisée pour décrire la dénaturation en chlorure de guanidinium. En effet la précision de nos mesures n'était pas suffisante pour faire la distinction entre plusieurs états, tous relativement déstructurés, et l'état intermédiaire que nous avons détecté peut en fait représenter une moyenne sur tous les états d'une distribution.

Il est très important de souligner que, pour caractériser ces états intermédiaires, nous avons réalisé des d'analyses différentes pour la dénaturation chimique et pour la dénaturation thermique. Pour pouvoir faire une comparaison précise entre les deux types de dénaturation, il faudrait d'une part réaliser des expériences d'ANS en température et d'autre part conduire une analyse SVD sur les courbes de diffusion obtenues pour différentes concentration en agent dénaturant.

Enfin, le schéma de dénaturation, que nous proposons pour la protéine dénaturée en chlorure de guanidinium et en température, s'apparente à celui proposé par Griko et al (Griko1994) à partir d'une étude thermodynamique sur les états de dépliement de la α -lactalbumine. Ces auteurs décrivent alors un model de hiérarchie coopérative (HC), la dénaturation thermique de la structure native s'effectuent en deux étapes : une première qui est très coopérative et à deux états, suivie d'un dépliement graduel de la protéine dénaturée, non coopératif. L'état dénaturé est décrit ici comme une distribution de divers sous-états dont le degré structural varie en fonction de la température.

En général il est possible que le dépliement puisse comporter une transition coopérative suivie par un graduel dépliement de la protéine, dû au fait que les états qui existent juste après la transition ne sont pas complètement déstructurés, mais plutôt compacts avec un cœur hydrophobe et des structures secondaires. Le degré de structure de ces interactions résiduelles dépend de la température mais aussi des conditions de solvant.